ERIC Document Resume	See Instruction Sheet	DATE OF RESUME	DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER	2. ORIGINATOR CONTROL NO(S).	3. IDENTIFICATION NO.	DO NOT WRITE IN SPACE BELOW
ED002107	N.A.	N.A.	

4. AUTHOR(S): Cate, Clarence C. and others

5. TITLE OF DOCUMENT ABSTRACTED: Probability and Statistics

6. PUBLICATION TITLE: Independent Learning Booklet

7. EDITOR(S): Mann, William

8. PLACE OF ORIGIN San Diego County Dept. of Education
ORG. 6401 Linda Vista
ST. San Diego, Calif.
CITY, STATE
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE Dept. of Education San Diego County, San Diego, California

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	1961	72

MICROFILM COPY AVAILABLE: YES [X] NO [ ]

COPYRIGHT GRANTED: YES [X] NO [ ]

COPYRIGHT RELEASE FILED: YES [X] NO [ ]

TYPE OF REPORT: PROPOSAL [ ] IN-PROGRESS [ ] FINAL [X] FOLLOW-UP [ ]

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR: Study Guide

16. INDEXING TERMS (6-12)

arithmetic curriculum	advanced students	independent learning	home study
elementary education	experimental approaches	student interests	mathematics
statistics probability	independent study		

17. ABSTRACT (MAX. 250 WORDS)

A booklet designed for elementary school students to be used independently from and in addition to the regular classroom curriculum in mathematics is presented. Elementary uses and concepts of probability and statistics are examined. Permutation, combination, and predictability are presented in the section on probability. The meaning of statistics, the concepts of data and grouping, measures of central tendency, and the concept of range are given in the section on statistics. Both sections include practical applications and problems. The teacher should identify potential users, orient them to the areas to be explored, and stimulate and maintain their interest in sustained independent effort.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION	20. DIVISION OR DEPT.	21. CONTRACT NO. / PROJECT NO.
	22. TOTAL COST OF PROJECT	
	23. DURATION OR PROJECT	U.S. Department of Health, Education, and Welfare Office of Education

INDEPENDENT LEARNING BOOKLET

ED002107

PROBABILITY AND STATISTICS

DEPARTMENT OF EDUCATION • SAN DIEGO COUNTY

REPRINTED BY PERMISSION
U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN PRINTED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED002107

INDEPENDENT LEARNING BOOKLET

Probability and Statistics

Writing Staff

CLARENCE C. CATE
ROBERT S. INGRUM
JANET S. BRIGGS

WILLIAM MANN, *Editor*

Published by
DEPARTMENT OF EDUCATION, SAN DIEGO COUNTY
6401 Linda Vista Road
San Diego 11, California

Copyright, 1961, by
Department of Education, San Diego County
Printed in U. S. A.

TO THE TEACHER

This Independent Learning booklet is designed to help you open new avenues of learning for students who show a special interest in mathematics and a readiness to study on their own. Such students will be able to work directly from printed materials, using abstract ideas to follow a line of mathematical reasoning. Occasionally, of course, even the ablest of them will need your help to get over some of the rough spots, but your main concern in connection with the booklet will be to identify students who can benefit from using it, to orient them to the area to be explored, and to stimulate and maintain their interest in sustained independent effort.

Since the topic of the booklet trends outward from the regular curriculum, the challenge to do this independent work should arise from students' own searching in the field or from new interests that you have succeeded in awakening. Students will profit most if they elect voluntarily to do the work rather than receiving it as an assignment or as a means of earning extra credit.

The material is so organized that students may progress at their own best pace. They should be encouraged to move ahead in discovering for themselves meanings and approaches in mathematics beyond those usually taught in the classroom, but there should be no requirement to complete parts or all of the booklet, or to follow a set schedule.

While serious effort will be necessary, students should find the work enjoyable because of their interest in the subject. They should gain a real sense of satisfaction through working individually or with other students in independently achieving understanding, overcoming temporary difficulties, and evaluating their own progress. Your encouragement and occasional direct help will of course be needed.

This booklet obviously is not aimed at developing complete understanding of an area of mathematics. Its purpose is rather to provide an introduction, or an approach novel to the student, which will create new interests and the desire to go on from here to further study in other sources.

A brief teacher's guide has been prepared to assist you in using this booklet with your students. Copies of the guide are distributed to your district office by the Department of Education, San Diego County. Additional copies are always available.

CONTENTS

part	topic	page
	Introduction	4
I	**PROBABILITY**	
	Mathematical and Statistical Probability	5
	How Do We Express Probabilities?	9
	How Do We Record Probability Experiments?	12
	Do Coins Have a Memory?	15
	What Happens When We Toss Several Coins?	18
	Six-Sided Probabilities	22
	How Many Banners? How Many Batters?	24
	How Can We Predict What Will Happen?	31
	Test Your Knowledge of Probability	35
	Checking Your Thinking	37
II	**STATISTICS**	
	The Meaning of Statistics	42
	Some Ways of Organizing Facts	43
	Measures of Central Tendency	48
	The Range	57
	Grouping Statistical Data	58
	Test Your Knowledge of Statistics	62
	Checking Your Thinking	64

PROBABILITY AND STATISTICS

INTRODUCTION

If you had your choice of three gifts of great value, would you choose wisely? Read the brief tale below — think — then choose.

> In ancient times, after a long journey three pilgrims arrived at an outer gate of Damascus. Tomorrow they would enter the city of their hopes; tonight they shared their last handful of food and prepared for sleep. At this moment they saw approaching a stranger whose manner and bearing told immediately that here was no ordinary being. The pilgrims listened with awe as the stranger offered to give them whatever they deemed of most value.
>
> The first pilgrim said, "O tall one-eyed stranger, give us money and jewels as much as we can carry, for as you see, we are the poorest of the poor!"
>
> Then the second pilgrim implored, "Let us become the Lords of Damascus that we might protect the old and devout, such as we."
>
> The third pilgrim asked, most humbly, "Let us see into our tomorrow, for we have expected much and long."
>
> Having heard these pleas, the mysterious stranger replied, "Of these I will give you the most valuable. . . ."

What do you think was the rest of the stranger's answer? Although there is no record of what he went on to say, perhaps you can decide after studying the modern mathematics of probability and statistics. In this study you will learn how we deal scientifically with situations that often seem to be matters of chance, coincidence, or luck. You will be introduced to ways of thinking about chance, about predicting the future, and about describing the past.

As you read the booklet, think carefully and be sure you understand each section before going on. Sharpen your thinking by completing each set of exercises. You will need pencil and paper to do some of the work, but please do not write in the booklet itself since other boys and girls will also use it.

Your teacher will help you if you have difficulty, but try to do the work on your own. A glossary at the end of the booklet defines some terms that may be unfamiliar to you.

PART I — PROBABILITY

Mathematical and Statistical Probability

It has often been said: "This is a world of chance." How can we deal with chance? How do we determine probabilities?

We might start with a simple illustration of tossing a coin to get heads or tails. We use coin tossing to illustrate probability because this allows us to assume "ideal conditions." What do we mean by "ideal conditions"? We mean that our coin is perfectly balanced (symmetrical) and always uniformly tossed so that there is no reason except chance for it to land with one side up rather than the other. The opposite of an ideal coin, for example, might be one that weighed more on one side and had a lump on the other so that its lack of symmetry would unavoidably affect the way it landed each time it was tossed.

Now, let us assume that we have ideal conditions, that we can distinguish one side from the other (tell which side is heads and which is tails), and are sure that the coin will always land with one of the sides up rather than on its edge. When the coin is tossed, how many possible ways can it land? Yes, 2; either heads up or tails up. Can it land both ways on one toss? Obviously not.

Since the coin must land heads or tails, but not both on the same toss, and since we have ideal conditions, how would you expect the coin to land? You would have no more reason to expect it to land heads than to land tails, so you would have to say, "It is just as likely for the coin to land tails as it is for it to land heads." For convenience, we will refer to **heads** and **tails** as the possible events. The event of landing on edge we ruled out, so this is not one of our possible events.

We can see that when we deal with this kind of probability there are two very important things to keep in mind. First, the two possible events (heads, tails) which can occur when we toss a coin are **mutually exclusive** because only one of them can occur at one time. When one occurs the other cannot occur on the same toss of the coin. Either heads must occur, or tails must occur, but both cannot occur together.

The second important thing to consider is that if a coin should be tossed under ideal conditions it is just as likely that heads will occur as that tails will. We call such events **equally likely.** In everyday speech, people say, "The chances are even," "It's a 50-50 proposition."

Now, what if you were asked to predict the results of tossing a ton of ideal coins? You could say without actually tossing the coins that the proportion of times the coins would land heads would be about equal to the proportion of times they would land tails. (The proportion of heads or tails is the number of heads or of tails divided by the total number of coins.)

You can make this statement about the probability of heads (or tails) because you assumed that you had ideal conditions and that the events were mutually exclusive as well as equally likely. You reached your conclusion by reasoning from what you already knew. You did not actually toss the coins, which would be experimenting, nor did you gather facts about the results other people might have gained in actually tossing coins. What you did was to determine the **mathematical probability** for the problem.

Mathematical probabilities, then, are found through reasoning alone and do not require experimentation. You should realize that there are occasions when, through studying the situation, you can determine a mathematical probability by using reason without going to the trouble of experimentation.

Determining Mathematical Probabilities	**Ideal conditions for the problem situation are defined.** **A mathematical model is designed.**	**Probabilities are determined by reasoning from the assumed conditions.**

Now we are going to look at another kind of probability. Suppose you decide to go ahead and actually toss a coin to determine the probabilities of heads and tails by experimenting. (Later you will have an opportunity to do this.) You will first be sure to make the conditions as close to ideal as you can so as not to favor one type of result over another. As you toss the coin

you will keep a record of your results. From this record you can determine what the proportion is of heads (or tails) to the total number of tosses. However, because you actually toss a coin and keep a record, you are experimenting and gathering statistics (as facts are sometimes called). Therefore, you are now determining the probabilities in a different way than you do when you determine mathematical probabilities. Such probabilities based on experience are called **statistical probabilities.**

Determining Statistical Probabilities	**Facts are collected through observation and experiment.**	**Probabilities are determined from facts selected as being pertinent to the situation.**
	Facts are arranged in statistical form.	

Although we distinguish between these two kinds of probability—mathematical probability and statistical probability—they are closely related. In statistical studies we use mathematics to organize our facts and to determine probabilities; on the other hand, part of our work in mathematical probability is to develop what we call **charts** or **models** to guide us in experimenting, gathering facts, and doing statistical studies.

We have talked about determining probabilities by actually tossing a coin; now let's look at a more complicated situation. How would you predict the probability of a light in your kitchen going on when you turn on the switch? As with tossing a coin once, there are just two possible events. Either the light will go on or it won't. Are these events **mutually exclusive**? Yes, they cannot both occur at the same time. But are they each equally likely?

No, we can't say the events are equally likely until we get additional information. We need to have such facts as the normal life of the switch as well as of the light bulb, how much each of them has been used, whether they are damaged in any way, whether the electricity is on, and so forth. Therefore, with what kind of probability are we dealing? Mathematical? Statistical? We know that the probability of the light going on is based on facts, or statistics. Thus, it is a statistical probability that can only be found by gathering facts.

A situation like this, having a variety of factors and lacking the strict controls we had in our experiment with coin tossing, illustrates why we usually need many facts to arrive at statistical probabilities on which we can depend. Remember, too, that when we say many facts we really mean many **related** facts, or facts pertinent to the situation. We must always think carefully and decide which statistics are needed. We then get as many of these pertinent facts as possible since our conclusion about the probability will be only as dependable as our facts.

Thus, in more complicated matters which are affected by many forces, we must gather facts or experiment to determine probabilities. However, in situations in which the factors are limited by previous definitions (such as the two ways a coin can land, or the distribution of cards in a deck of playing cards) it is unnecessary work to figure out statistical probabilities.

Stop now to check your understanding up to this point. Think about the questions in Set A; then check your thinking on page 37.

SET A

1. In tossing a coin, what do we mean by mutually exclusive events?
2. In tossing a coin, what do we mean by equally likely events?
3. What kind of probability is based on experience or experimentation?
4. What kind of probability is based on reasoning about a previously defined situation?
5. What kind of probability is involved in determining the following?
   a. The probability of drawing an ace on the first try from a standard deck of playing cards (to be determined without actually drawing any cards).
   b. The probability that a 30-year-old person will live to be 60.
   c. The probability that the weather forecast will be right.
   d. The probability that a fairly new drug will be successful when it is given to a laboratory animal.

How Do We Express Probabilities?

How do we find and express the chances or probability that a particular event, which we will call **the given event,** will occur? How do we show this as a numerical value? You have used the method employed for this purpose many times. Probabilities are shown as fractions, as **ratios.** A ratio is the numerical value of one quantity or amount divided by the numerical value of a second quantity.

What is an example of a probability shown as a ratio? We have seen that the chances that an ideal coin will land heads are one in two. We can express this relationship of "one chance in two" as $\frac{1}{2}$. Written as a formula we have:

$$\text{Probability of heads} = \frac{1}{2}$$

Here the fraction $\frac{1}{2}$ is a numerical ratio which describes the chances of a tossed coin coming to rest with the head side up. How do you show the chances that the coin will land tails? Yes, also $\frac{1}{2}$.

Look at this fraction: $\frac{1}{2}$. What does the 1 in the numerator stand for? What does the 2 in the denominator stand for? You can see that the numerator (1) stands for the number of ways one of the events can occur, while the denominator (2) stands for the total number of possible events (heads **and** tails). We can now write our complete formula in words:

$$\text{Probability of a given event} = \frac{\text{the number of ways a given event can occur}}{\text{the total number of possible events}}$$

We can change the common fraction $\frac{1}{2}$ to the decimal fraction .50, since we use both decimal fractions and common fractions to express probabilities. We can say that this probability is $\frac{1}{2}$ (one out of two), .50, or 50 percent.

We sometimes hear people say: "There is a 50-50 chance that it will happen." Though this is slang, it is an effective way to indicate an even chance. It means that out of 100 events, 50 events would be the given event while the other 50 would not.

We have been discussing the probability involved in tossing a coin which has 2 mutually exclusive **and** equally likely events. The mathematical probability is $\frac{1}{2}$, or .50. Let us now consider how this probability of .50 might compare with the probabilities which could be found in quite different situations. Stop to think. What **are** all possible probabilities? What are the extremes of probabilities? What is the **range**, or **difference**, between the extremes?

How probable can something be? Can it be a "must"? Consider: If you have only one pencil in your desk and it is yellow, and you open the desk to take out one pencil, what is the probability that you will take out a yellow pencil? The probability is **certain**, isn't it? The given event **must** occur.

How can we express this certain probability as a ratio? The number of ways the given event (your taking out a yellow pencil) can occur is 1. The number of possible events (your taking out a yellow pencil) is also 1. Substituting in our formula we get the fraction $\frac{1}{1}$ or 1. This can also be written 1.00.

Let's look at another certain event to be sure that we get the idea. When you toss an ideal coin you know it is certain to land with **either** heads or tails showing. It must be one of these but it cannot be both on one toss. Therefore the probability (P) that it will be **either** heads **or** tails is:

$$\text{P of either heads or tails} = \frac{\text{The number of ways the given event can occur—heads or tails}}{\text{The total number of possible events—heads and tails}}$$

By substituting numbers we get:

$$P = \frac{2}{2}, \text{ which } = \frac{1}{1}, \text{ which } = 1 \text{ or } 1.00$$

Another way of looking at this is to recall that the probability of heads on one toss is $\frac{1}{2}$ and the probability of tails on one toss is $\frac{1}{2}$. The probability that the coin will either land heads or land tails must be equal to the probability that it will be heads **plus** the probability that it will be tails. Add these two probabilities: $\frac{1}{2} + \frac{1}{2} = 1$ or 1.00. Does this agree with what we just saw—that 1.00 is the probability of an event that is certain to occur?

Now, how improbable can something be? Once again this extreme is obvious. The event can be impossible—something that can't occur. Let's go back to the example where you have only one pencil (yellow) in your desk. If you take one pencil out of your desk, what is the probability that it will be red? This event can't happen, can it? Written as a fraction, this probability becomes:

$$\text{P of taking out a red pencil} = \frac{\text{The number of ways the given event can occur}}{\text{The total number of possible events}}$$

By substituting the appropriate numbers for the words you get $\frac{0}{1}$, or 0, since zero divided by any number equals 0.

Thus probabilities can range from 0 to 1 (or from .00 to 1.00). Probabilities of events other than certain or impossible ones are found between these two extremes. As we saw, the probability of throwing heads on one toss of an ideal coin is .50. At what point in this range (.00 to 1.00) is .50 located? Yes, it is in the middle of the range.

In the previous section you became familiar with the difference between mathematical and statistical probabilities. In this section you looked at how to express probabilities. Try to apply what you have learned as you do the examples in Set B. Express your probabilities as ratios or percents. Check your thinking by turning to page 37.

SET B

1. What is the probability that someone now alive will be dead 200 years from now?
2. What is the probability that water will freeze at 50° F.?
3. Joe puts 12 red balls in an empty box. He takes out one.
   a. What is the probability that he takes out a black ball?
   b. What is the probability that he takes out a red ball?
4. List some events that have probabilities of 1.00.
5. List some events that have probabilities of .00.
6. What is the probability that you are 6 years of age?
7. What is the probability that you are younger than your father?

How Do We Record Probability Experiments?

We said that we would toss a coin to determine the statistical probability of mutually exclusive, equally likely events. To do this let us make a list or table of all the possible events that can occur. How many will this be if you are going to toss only one coin? Yes, 2—heads and tails.

Make a table like the one below. We call this a "tally sheet."

Heads	
Tails	

Now get a fairly new penny. Be sure it is not bent or worn. Toss the penny into the air and let it fall on a smooth level surface. Note whether the coin falls heads up or tails up, then make a mark, or "tally," after the right word on your tally sheet to record which way it fell. When you have four tally marks in a row like this, //// , cross them with the fifth mark like this, 𝍸 . This groups each five tally marks together so that they are easy to count. Now you are ready to make some tests. All of your answers to Set C except Example 9 will depend on the results of your coin tosses. You can check the answer to Example 9 by turning to page 38.

SET C

1. Toss the penny 10 times and keep a record on your tally sheet. How many heads were there? How many tails? How many heads and tails combined?
2. Toss the penny 20 times. Were there two times as many heads as there were when you tossed it 10 times? How many heads were there? How many tails? How many heads and tails combined?

3. Count all the tally marks for Examples 1 and 2. How many heads were there in 30 tosses? Write the number that tells the chances of getting heads according to the results of these 30 tosses.

4. Toss the penny 20 more times. Count the tally marks for all 50 tosses. Did you get a more nearly equal number of heads and tails in 50 tosses than you did in 10 tosses?

5. Write the probability ratio (fraction) for tails from the 10 tosses in Example 1. Remember that the denominator will be the total number of events (10); the numerator will be the number of the given event (tails). Then write the probability ratio for heads in Example 1.

6. Write the probability ratio for tails from the 20 tosses in Example 2. Write the probability ratio for heads in Example 2.

7. Write the probability ratio for tails from the 30 tosses in Example 3. Write the probability ratio for heads in Example 3.

8. Write the probability ratio for tails for the 50 tosses in Example 4. Write the probability ratio for heads for the 50 tosses in Example 4.

9. The two ratios in Example 8 above are common fractions. Find their sum.

Now that you know how to record the results of coin tossing experiments, you are ready to give some attention to recording experiments which are more complicated. For example, how would you go about determining the following probability?

The chances that it will rain on the day of the class picnic.

Look at this example. What must you ask yourself first? Yes, you need to know how many possibilities there are. How many are there? Yes, two. Are they mutually exclusive? Yes, they both can't happen at once. Either it will rain that day or it won't. Your next question is: Are these chances equally likely? If they are, the probability is mathematical and you can determine it at once. That is, the chance that one of them will happen must equal the chance that the other one will; each chance then would be one out of two, or $\frac{1}{2}$.

But you can't know beforehand that these chances are equally likely, because many different forces determine changes in the weather. You have a complicated problem in statistical probability that requires the gathering of many facts and statistics (data).

How would you go about gathering and listing your data? Could you use a tally sheet similar to the one you used for tossing the coin?

Before you answer this question, stop to think what kind of data you will need. Does this example remind you of the example about your kitchen light's chances of going on? Many different kinds of information were needed to determine that probability, weren't there?

Your method of recording information for such kinds of prediction will certainly be more involved than the tally sheet system you used for the coin tossing experiment. For example, you may want to keep a daily record of rainfall for several months before the picnic. You may also want to get the rainfall records for the date of your picnic for past years. You will record such data in tables like those in Part II of this booklet.

Though you can't be certain whether it will rain or not, your estimated probability will be more dependable if you have gathered many statistics and evaluated them properly.

Do Coins Have a Memory?

Let us say that you toss an ideal coin once. It comes up heads. You toss it again. Heads, once more. You toss it a third time. Heads again. You want to toss it a fourth time. What are the chances that it will be heads this time?

To answer this question let's go back and see what the chances were the first time you tossed the coin. Each possibility, heads or tails, was equally likely, wasn't it? Therefore the chances were one out of two, $\frac{1}{2}$, or .50 that it would land tails.

What about the chances on the second toss? Yes, each possibility was still equally likely. The coin couldn't remember that it came up heads the first time, nor could it conclude that it therefore should come up tails the second time. It landed the way it did by chance and not because it chose to land one way or the other. The chances were still $\frac{1}{2}$ or .50 the second time that it would land heads.

What were the chances on the third toss? Yes, still $\frac{1}{2}$ or .50 for heads, and the same for tails. By using this kind of reasoning you can see that **each time** you toss an ideal coin the chances are $\frac{1}{2}$ or .50 that it will land heads and $\frac{1}{2}$ or .50 that it will land tails, regardless of what it landed the last time, the last three times, or the last hundred times! Each toss of the coin is completely independent of the previous tosses. Thus the chances that our coin will land heads the fourth time are still $\frac{1}{2}$ or .50.

How does this principle work if many tosses are made? Is it correct to say that if a coin is tossed 10,000 times it **ought** to land heads as often as it does tails? No, it is not correct to say this unless we mean that the proportion or percentage of heads out of 10,000 tosses will approximate mathematical probability.

Our actual results from 9,999 tosses might be 4,890 heads and 5,109 tails. Do you see that $\frac{4890}{9999}$ is approximately $\frac{49}{100}$, or .49, and $\frac{5109}{9999}$ is approximately $\frac{51}{100}$, or .51? Each is close to .50 or $\frac{1}{2}$. But, while the proportion of heads (4890/9999) is about equal to the proportion of tails (5109/9999), there are considerably fewer heads than tails. Can we conclude that because

there are fewer heads the ten-thousandth toss **ought** to be heads or that the chances are better for heads than they are for tails? No. The chances are still the same on the last toss of even as many as 10,000 tosses: $\frac{1}{2}$ for heads and $\frac{1}{2}$ for tails.

If you get this big idea that the probability of equally likely events is not changed by past results you have grasped an idea which many people have missed. Even famous mathematicians have fallen into the trap of believing that the one of two equally likely events which has happened less often should begin to happen more often in order to catch up with the other.

People also make the mistake of believing that because one of two equally likely events has happened more often than the other it will continue to do so. Will it? By definition, the chances for equally likely events to occur remain the same for each separate occurrence. Therefore you should not expect one of the events to continue to run ahead, to run behind, or to catch up as you enter marks in your tally sheet.

We have just illustrated the principle of independent events. We can state it this way: Two events are independent if the occurrence or non-occurrence of one is not affected by the occurrence or non-occurrence of the other. We can show the principle by tossing a coin twice; the event heads on the first toss will have no effect on whether you get heads or get tails on the second toss.

Stop to think about the number of tosses required to illustrate the principle of independent events in order to avoid confusing the idea of "independent events" with that of "mutually exclusive" events. We illustrated mutually exclusive events by a **single** toss of coin, saying that the event heads prevents the event tails—you can get either heads or tails, but not both. We illustrate independent events by **two or more** tosses of a coin, saying that the result of the first toss will have no effect on the result of the second toss, and so on.

What does this principle of independent events which are equally likely mean to us? Why are we concerned with it? As you know, we have problems or situations involving questions of chance every day. To know how to solve these problems we have to figure out what kind of chances we are dealing with.

We must really understand the principle of events which are equally likely so that we can identify such events when we run across them. If we know that the possibilities are really equally likely we will want to determine mathematical probabilities. We will not have to experiment and gather statistics. However, if we are not sure that the possibilities are equally likely or if we know that they are not, we will have to experiment and get statistical probabilities.

Now answer the questions in Set D below. Check your thinking by turning to page 38.

SET D

1. What are the chances that an ideal coin which has landed tails four times in a row will land tails the fifth time?
2. An ideal cube has 6 flat surfaces. Each of these is identical except that each is painted a different color: red, green, blue, black, yellow, and orange. If you toss the cube one time, what are the chances that the red surface will come up?
3. We tossed our ideal cube described in Example 2 several more times. It came up black the first time, yellow the second, green the third, blue the fourth, and red the fifth.
   a. What are the chances that it will come up orange the fifth time?
   b. What are the chances that it will come up black the sixth time?
4. We tossed the cube again. We got a black, a yellow, a red, and a blue out of 4 tosses.
   a. What are the chances that either green or orange will come up the fifth time?
   b. What are the chances that either black or yellow will come up the fifth time?
   c. What are the chances that the sixth toss will come up either red, blue, or yellow?
   d. What are the chances that the sixth toss will come up one of these: black, yellow, red, blue, green, orange?
   e. What are the chances that purple will come up on the sixth toss?

What Happens When We Toss Several Coins?

If we toss two coins together, for example 2 pennies, four possible events can occur. What are these events? Each event includes both coins, since we are tossing them both at the same time. The model below shows us what these events are.

MODEL FOR TWO-COIN EXPERIMENT

Both may fall heads (HH)
The first one may fall heads and the other tails (HT)
The first one may fall tails and the other heads (TH)
Both may fall tails (TT)

Coins	Possible Events			
First Coin	H	H	T	T
Second Coin	H	T	H	T

H = Heads
T = Tails

Each column shows an event—2 coins tossed.

Since there are four possible events, what is the probability that any particular one of them will occur? (We shall call this event **the given event.**) Use your formula to figure out this probability.

$$\text{Probability of the given event} = \frac{\text{The number of ways the given event can occur}}{\text{The total number of possible events}}$$

Let's figure the probability that both of the coins will fall heads. How many ways can this event happen? Yes, 1 (as shown in the first column). What is the total number of possible events? Yes, 4 (first, second, third, and fourth columns). Therefore the probability for both coins to fall heads is $\frac{1}{4}$. What is the probability that both coins will be tails? You can see that it also will be $\frac{1}{4}$. What is the probability that the first coin will be heads and the second tails? It is $\frac{1}{4}$ again. This same probability would hold for the first coin to be tails and the second to be heads.

Now think: What is the probability for the given event that **either** coin will be heads and the other tails? Is it $\frac{1}{4}$ again? As you answer, look carefully at your model for the two-coin experiment. Think about how we described the given event (either coin heads and the other tails). How many ways can this given event occur? There are 2 possible ways, since we don't care which of the coins is heads or which is tails. We say only that either can be heads and the other tails. Now, use your formula to determine the probability. You can see that it is $\frac{2}{4}$, or $\frac{1}{2}$, or .50.

Suppose now that we toss three coins at once. The model below shows all the possible events that can occur.

MODEL FOR THREE-COIN EXPERIMENT

Coins	Possible Events							
First Coin	H	H	H	T	H	T	T	T
Second Coin	H	H	T	H	T	H	T	T
Third Coin	H	T	H	H	T	T	H	T

Each column shows an event—3 coins tossed.

How many possible events are there? What is the probability that any one of the events will occur? Check your model, then use the formula again.

Yes, the probability of any one event is $\frac{1}{8}$. Now let's think about the probability of this type of given event: Any 2 of the coins can be heads and the other tails. Is this probability also $\frac{1}{8}$? Before you answer, look again at your model for the three-coin experiment. In how many ways can this event occur? Yes, there are 3 possible ways since we don't care which two of the coins are heads or which one is tails. Use your formula again to determine the probability. What do you get? Yes, $\frac{3}{8}$. What is the probability that any 2 of the coins will be tails and the other heads. Check your model. In how many ways can this event occur? Yes; P $= \frac{3}{8}$. Let's look at the chart on the following page which shows the ways three coins can fall.

CHART FOR THREE COINS

Possible Ways for Three Coins To Fall	Number of Times Each Way Can Occur
3 Heads	1
2 Heads and 1 Tail	3
2 Tails and 1 Head	3
3 Tails	1
Total Events	8

You can see by looking at the chart that the probability of 3 heads is $\frac{1}{8}$. What is the probability of 3 tails? Yes, also $\frac{1}{8}$. What is the probability of 2 heads and 1 tail? Yes $\frac{3}{8}$. Of 2 tails and 1 head? $\frac{3}{8}$ again. Now what is the probability that 1 of the coins will not be the same as the other 2? Can you see that it would be the sum of the probabilities for 2 heads and 1 tail and 2 tails and 1 head? You could also get this probability by subtracting the sum of the probabilities that the coins would be all the same (all heads, $\frac{1}{8}$, and all tails, $\frac{1}{8}$) from the total possibilities ($\frac{8}{8}$ or 1). In either case the probability that one of the coins will not be the same as the other two is $\frac{6}{8}$ or $\frac{3}{4}$.

The four possible events shown by our chart which occur when we toss 3 coins are mutually exclusive. As we said, this term means that when one of them occurs, the others cannot occur. We learned earlier too that we can add the probabilities of mutually exclusive events when we want to find the probability that any one of several such events can happen.

We said that the sum of the probabilities of all these events is 1. We can prove this by adding these probabilities.

3 Heads	$\frac{1}{8}$
2 Heads and 1 Tail	$\frac{3}{8}$
2 Tails and 1 Head	$\frac{3}{8}$
3 Tails	$\frac{1}{8}$
	$\frac{8}{8} = 1$

If we want to know the probability of getting either 3 heads or 3 tails we add $\frac{1}{8}$ and $\frac{1}{8}$ and get $\frac{2}{8}$ or $\frac{1}{4}$. If we want to know the probability of getting 2 or more heads on one toss, we add $\frac{1}{8}$ (for 3 heads) and $\frac{3}{8}$ (for 2 heads and 1 tail) and get $\frac{4}{8}$ or $\frac{1}{2}$.

We can also state the probability that one of these events will **not** occur. Since the sum of the probabilities for all the possible events is 1, we can find the probability that 3 heads will not occur by subtracting $\frac{1}{8}$ from 1. ($P = \frac{7}{8}$.) You can prove this by counting the number of events other than 3 heads which can occur when 3 coins are tossed.

Now try the examples in Set E. Check your thinking by turning to page 38.

SET E

1. If three ideal coins are tossed, what is the probability that one of them will be heads?
2. Two ideal coins are tossed:
   a. What is the probability that both will be tails?
   b. What is the probability that neither coin will be tails?
   c. What is the probability that at least one coin will be heads?
3. If three ideal coins are tossed, what is the probability that none of them will be tails?

The chart below shows the possible events when 4 coins are tossed. Use it to answer the questions below it.

Possible Events	Number of Ways Each Event Can Occur
4 Heads	1
3 Heads 1 Tail	4
2 Heads 2 Tails	6
3 Tails 1 Head	4
4 Tails	1

SET F

1. In how many ways can all of the possible events occur when 4 coins are tossed?
2. What is the probability of 4 heads?
3. What is the probability of 2 or more heads?
4. What is the probability of 3 or more tails?
5. What is the probability of an equal number of heads and tails?

Check your thinking on page 38.

Six-Sided Probabilities

Now that we have worked with several coins at once, let's look at another kind of probability situation which requires us to choose from many possible events. Perhaps you have used cubes in games like Parchesi, Monopoly, Finance, etc. In such games the players must travel from start to finish over a path divided into spaces or steps. Each player has the same number of spaces to travel. The player who gets to the finish line first is the winner.

What is the simplest way for us to make sure that every player has an even chance to win? We need something which will tell a player how many spaces to travel each time he gets a turn. For this purpose, a die (or dice if more than one) is best because of the way it is made. Each die has 1, 2, 3, 4, 5, or 6 dots on each of its sides or faces. When a die is thrown, it may come to rest with any number of dots, from 1 to 6, on top. The number of dots on top are the ones that count, or the ones you go by. You can see that the probability of getting any one of the numbers is $\frac{1}{6}$ because there is an equal chance that any one of the six will be thrown.

When two dice are thrown there are then two sets of dots, one set on each die. This model shows all the possible sums of these pairs of numbers.

MODEL FOR SUMS OF TWO DICE

SECOND DIE \ FIRST DIE	1 dot	2 dots	3 dots	4 dots	5 dots	6 dots
1 dot	2	3	4	5	6	7
2 dots	3	4	5	6	7	8
3 dots	4	5	6	7	8	9
4 dots	5	6	7	8	9	10
5 dots	6	7	8	9	10	11
6 dots	7	8	9	10	11	12

Count the sums in the model to find the number of possible events. You can see that most sums occur in more than one way. Study the model; then get paper and pencil and complete the table which has been started for you below. Then do the examples in SET G. Check your answers on page 38.

Sum of Dots on Two Dice	Number of Ways This Sum Can Occur
2	1
3	2
4	3
5	4
6	5
7	6
8	5

SET G

1. What sums are least likely to be thrown with 2 dice?
2. What is the probability of getting a sum of 7 with 2 dice?
3. What is the probability of getting a sum of 5 with 2 dice?
4. What is the probability of getting a sum smaller than 5 with 2 dice?
5. What is the probability of getting a sum other than 7 with 2 dice?
6. What is the probability of getting an even sum with 2 dice?
7. What is the probability of getting an odd sum with 2 dice?
8. What is the probability of getting a sum larger than 4 but smaller than 10 with 2 dice?
9. What is the probability of getting a sum of 6, 8, or 10 with 2 dice?
10. What is the probability of getting a sum of 11 with 2 dice?
11. What is the probability of getting a sum of either 7 or 11 with 2 dice?
12. What is the probability of not getting a sum of either 7 or 11 with 2 dice?
13. What are the chances of getting a sum of 1 with 2 dice?
14. What is the probability that each of the 2 dice will show an odd number of dots?
15. What is the probability that each of the 2 dice will show the same number of dots?

How Many Banners? How Many Batters?

Suppose you and your classmates are going to make a school banner. You decide to make it of 3 different colors. The colors you select are green, white, and red. In how many different ways can you arrange these colors?

You might find out the number of different ways possible by making a "tree" diagram like the following. (Can you see why this is called a "tree" diagram? Think of how a tree stands with a trunk, then large branches, and then smaller branches.)

Left Side	Middle	Right Side	Arrangements
Green	White	Red	1
	Red	White	2
White	Green	Red	3
	Red	Green	4
Red	Green	White	5
	White	Green	6

By looking at your tree diagram you can see that there are 6 different ways to arrange the three colors. Do you see why you can get six banners out of only three colors? Yes, it's because each time you change the **order** of the colors you get a new **arrangement** of colors and therefore a new banner.

The tree diagram is valuable in listing all the possible arrangements you can make. It helps you to be sure that you have not missed any possibility, or used any arrangement more than once.

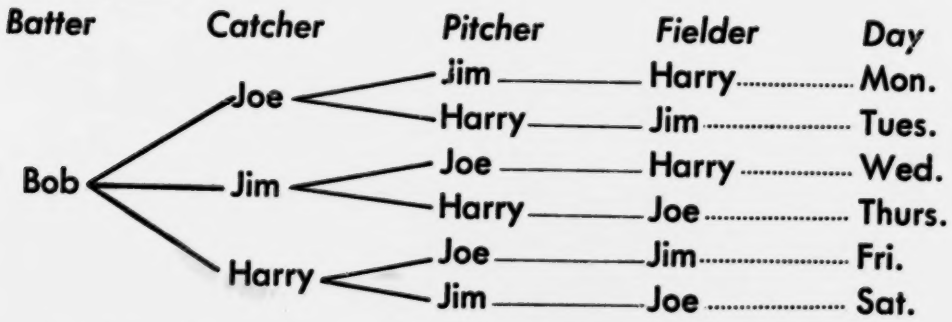
Now let's look at a slightly more complicated example to see how we use a tree diagram. Perhaps you have played games which require that the players take turns. You have to decide

who will be first, who will be second, and so on. In other words, you have to establish the **order** of all the players in the game. Let's take a look at how Bob and his three friends went about doing this.

Bob and his friends Joe, Jim, and Harry meet every afternoon during the baseball season to play work-up, a kind of softball which can be played by four or more players. All the boys like to be batter. The game is played by a batter, a catcher who is next up to bat, a pitcher who is up third, and a fielder who is fourth to bat. The boys were trying to decide how they could line up for their afternoon games so that during the season everybody would get a fair share of chances to be first at bat.

Their coach helped them write down all the different orders in which the four boys could line up. Knowing this, they could use a different line-up each day. The coach started by showing them how to make a tree diagram using one boy, Bob, as the first batter. It looked like this:

Diagram A:

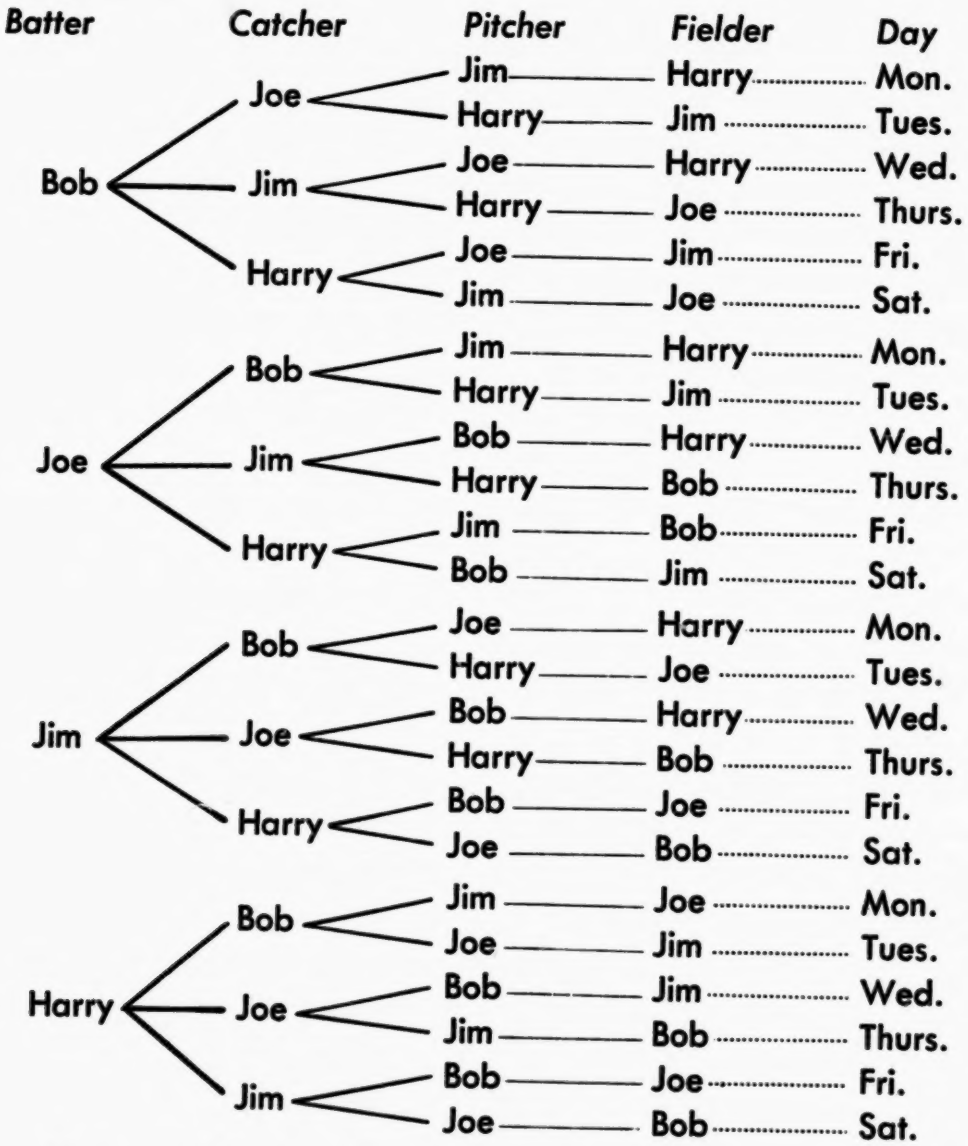


On Monday Joe would be catcher, Jim pitcher, and Harry fielder; on Tuesday Harry would move up to be pitcher and Jim would be fielder. You can see from the diagram how the boys shifted during the rest of the week.

The boys learned that when they selected one boy as batter, any of the other three boys could be named as catcher for the first day, either of the remaining two boys could be selected as pitcher, and the fourth or last boy would be the fielder. One boy would be the first batter for six days in a row, while the rest of the boys would have different positions from day to day. To make their schedule fair for all of them, the four boys had

to expand their tree diagram to cover four 6-day periods of playing. In this way each boy could be first batter for a 6-day period. By studying the diagram that the coach had helped them draw for the first week, the boys were able to complete a schedule for four weeks of playing. We show their tree diagram of this schedule below.

Diagram B:



Can you imagine how surprised the boys were to find that there were 24 different line-ups for a simple game!

Now, before you go on, you may like to try figuring out how many "line-ups" you can make, using the names of some of your own friends. Drawing tree diagrams will help. Start

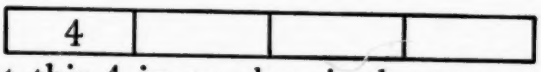
with three friends, next try it with four friends. Then study the next few paragraphs to learn a quick way of determining the number of possible ways of arranging colors, people, or things.

Bob studied the tree diagram that he and his friends had made for work-up. He remembered how surprised they were to find out how many different ways they could line up with four players. After a while he said, "I think I see a short cut for finding the number of ways we can line up. If we only want to know the total **number** of ways, we can find the answer without making a tree." Let's see how Bob was thinking.

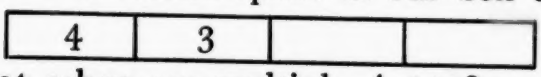
"We chose one of the **4** boys to be first batter, next chose any one of the **3** boys left to be catcher, then chose one of the remaining **2** boys to be pitcher, which left only **1** boy to be the fielder. If I multiply $4 \times 3 \times 2 \times 1$, I get 24, which is the same number of arrangements we got by counting with our tree diagram."

Bob talked this over with his teacher, who showed him how to make a box diagram like the one below. It is much simpler to use than a tree if you want to arrive at an answer quickly. Because we have 4 boys to arrange in all possible orders, we make our box diagram with four spaces.

In the first space we put the number of possible first choices for batter.

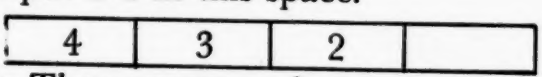


(Note that this 4 in our box is the same as the number of boys in the first column of our tree diagram.) Now for each of our second choices (for catcher), we have 3 boys, so we put this number in the second space in our box diagram.



Notice that when we multiply $4 \times 3 = 12$, we get the same number as when we count the boys in the column headed "catcher."

Now, in how many ways can the third space be filled? Yes, we have only 2 boys remaining for each of our choices for pitcher, so we put a 2 in this space.



Now multiply. Then compare the product of these three numbers with the count of the boys in the pitcher's column. Are

they the same? In the last space we put the number 1, which represents the one remaining choice, fielder. Note that for each of the 24 ways of filling the first three spaces we have one way of filling the fourth space in each case, because only 1 boy is left for fielder.

4	3	2	1

Multiply and you see that we get the same total product, 24, that we do when we count the number of boys in the fourth (fielder's) column.

You can see that using a tree diagram helps us to see all the different ways, or orders, in which we can arrange a group of people, colors, or really any kind of things. The arrangement of a group, or set, of things in a definite order or line is called a **permutation.** What we have been doing is to find how many permutations or different orders or arrangements we can make with the objects in a group, or set.

Up to now we have worked with permutations involving the same kinds of things—in the one case all were colors, in the other all were boys. Let's take a look now at an example in which we have to find various arrangements of **different** kinds of things.

Bob and Joe meet Mary, Susan, and Jane at the entrance to the school cafeteria. We have three questions concerning the way they might line up.

(1) In how many different ways can they line up to enter the cafeteria? (Because we ask only in how **many** different ways, and do not need to know specifically what each of these ways is, you can use a box rather than a tree diagram, although if you wish to make a complete check you can use both the box and the tree.)

(2) In how many different ways can the children line up if one of the boys is first? (A boy **or** a girl can be second.)

(3) After you have decided question No. 2, ask yourself: In how many different ways can they line up if one of the girls is first, and if either a girl or a boy is second?

Take a minute to go back over the questions. Then study the explanation on the next page.

(1) If any arrangements of the five youngsters can be made, you can figure the same way that Bob did in figuring out the permutation of four boys. What number would go in the first space this time? Yes, this time you would start with 5 (since there are 5 people), and your box diagram would look like this:

5	4	3	2	1

How many permutations are possible? Multiply to find out. Did you get 120? Would you have thought that there were so many?

(2) But if a boy is first, and either a boy **or** a girl can be second, your reasoning will be like this: Since there are 2 boys, I put a 2 in the first-choice space. Then I put the number of boys **and** girls remaining (4) in the second space.

2	4			

Continue, as you now know how to do, by filling in the spaces. Your box will look like this:

2	4	3	2	1

How many permutations have we found if a boy is first? The product is 48. If you make a tree, you will find that this product agrees with your count.

(3) How many permutations are there if a girl is first? Since there are 3 girls, the number in the first space must be a 3. Either a boy or a girl can be second, and there are 4 youngsters remaining, so 4 will be in the second space. Fill in the remaining spaces as usual.

3	4	3	2	1

How many permutations are there if a girl is first? Yes, 72.

We have found that there are 48 possible arrangements if a boy is first, and 72 possible arrangements if a girl is first. What happens if we add these? Shouldn't we get the same total number of possible arrangements that we had when we answered our first question when **either** a boy or girl could be first? Do you get the same total?

This idea of permutation is used in many ways. One common use we see every day is in license numbers on automobiles.

Suppose you wish to know how many different license plates you can make using the letters A B C and the digits 1, 2, 3. In how many ways can you arrange the 3 letters, A B C, if you use each letter only once in each permutation? Use Bob's method of multiplying and you find that the answer is $3 \times 2 \times 1$, or 6 ways. The same reasoning applies to the three digits.

What happens, however, if we use each letter more than once? Then we can choose any of the 3 letters for first place, any of the 3 letters for second place, and any of the 3 letters for third place. For instance, we are allowed to use A for first place, also for second place, and again for third place, and so on with all three letters. Again use the multiplication method and you find that the answer is $3 \times 3 \times 3$, or 27 arrangements of the three letters.

Now do the examples in Set H. Check your thinking on page 39.

SET H

1. How many permutations of the digits 1, 2, and 3 can you make if you use each numeral once in each permutation?
2. How many permutations of the three digits in Example 1 are possible if you use each digit more than once in any permutation?
3. How many license plates can you make by using the letters A B C and the digits 1, 2, 3 if each license plate has 3 letters and 3 digits? Each letter and each digit may be used more than once. For instance, one license number could be AAA 111.
4. Find the number of permutations of the letters in the word **g r a d e s** . Each letter can be used only once.
5. California uses 3 letters and 3 digits on each vehicle license plate. For instance, one California license number is TMY 531. Although there are 26 letters in the alphabet, California uses only 24 of them, omitting O and I since these letters might be confused with digits. How many different license plates can California have using the 24 letters and 10 digits?

How Can We Predict What Will Happen?

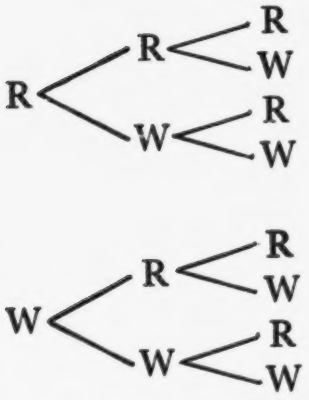
When we know the total number of different events that can occur or are possible, then we can predict how likely it is that any one of these events will occur. For example, if you have one red, one white, and one blue marble in your pocket and you take out just one marble, what are the chances that it will be the red one? Since there are 3 marbles, only one of which is red, it is easy to see that there can be only one chance in three that the first one you remove will be red. So we say that the probability of the given event—"red is the first marble"—is 1 in 3, or $\frac{1}{3}$.

If, however, the first marble you take out is a blue one, what is the probability of taking out the red marble next? In other words, what is the probability that the second marble will be red? Since only two marbles remain, the probability of choosing red is now 1 in 2, or $\frac{1}{2}$. If the first marble removed is blue and the second marble is white, then only the red one is left. In this case, of course, the probability of removing the red one the third time is now 1 in 1, or $\frac{1}{1}$, or 1.00, because this is the only possible way in which the given event can occur with only one marble (red) remaining.

As you have learned, the probability of any one of a set of events occurring is usually written as a ratio or common fraction in which the denominator is the total number of events that can occur and the numerator is the number of ways in which the given event, or the event we want to predict, can occur. If you have two red marbles and one blue marble in your pocket, the probability of choosing a blue one first is 1 in 3, or $\frac{1}{3}$. However, since there are two red marbles, the probability of choosing a red one first is 2 in 3, or $\frac{2}{3}$. Remember that the sum of all the probabilities ($\frac{1}{3} + \frac{2}{3}$) is always equal to 1. Now do the examples in Set I. Check your work by turning to page 39.

SET I

1. Jack's teacher has 12 pencils to hand out to his group. Of the 12 pencils, 4 are red and 8 are yellow. The teacher gives Jack one pencil.
   a. What is the probability that it is red?
   b. What is the probability that it is yellow?
2. The first five children receive yellow pencils.
   a. What is the probability that the sixth child will get a yellow one?
   b. What is the probability that the sixth child will get a red one?
3. Joe has 3 white and 3 black socks in his dresser drawer.
   a. How many socks must he take out to be sure that he has two of the same color?
   b. What is the probability that the first sock Joe takes will be a white one?
   c. What is the probability that it will be a black one?
4. Suppose that the first sock Joe takes out is white.
   a. What is the probability that the second one will be white also?
   b. What is the probability that the second one will be black?



Bob has 3 red and 3 white marbles in his pocket. If he wishes to take out 3 of these 6 marbles, the tree diagram at the left shows all the possible orders in which they can be removed. From this diagram, we can write these orders as a table.

Red Red Red	White Red Red
Red Red White	White Red White
Red White Red	White White Red
Red White White	White White White

You can see that there are 8 different permutations, or orders, in which Bob can remove 3 of these 6 marbles. Thus,

there are 8 possible events that can occur. We can find the probability of any of these happenings by counting the number of times it can occur in the 8 possible events. For instance, if we wish to know the probability of Bob's removing 3 red marbles we can see that this can happen only once in the 8 possible events. Therefore, the probability that he will remove 3 red marbles is 1 in 8, or $\frac{1}{8}$. What is the probability that the first two marbles he removes will be red? Yes, this probability is 2 in 8, or $\frac{2}{8}$, or $\frac{1}{4}$.

Now try the examples in Set J. Use the table to help you answer them. Then do SET K. Check your thinking by turning to page 39.

SET J

1. a. What is the probability that the first marble Bob removes from his pocket will be red?
   b. What is the probability that it will be white?
2. What is the probability that the first marble will be red and the second marble will be white?
3. What is the probability that the first marble will be red and the third marble will be white?
4. What is the probability that if Bob takes out 3 marbles, none of them will be red?
5. If the first marble Bob removes is white and the second is red, what is the probability that the third will be red? The correct answer is 1 in 2, or $\frac{1}{2}$. Can you explain why?

SET K

1. Two of the 8 boys in Bob's Scout Troop are going on a hike.
   a. What is the probability that Bob is going on the hike?
   b. What is the probability that Bob will not go on the hike?
2. Bob has two brothers. What is the probability that he is the youngest.
3. There are 32 students in Mary's class. What is the probability that Mary is the youngest child in the class?

Now, just for fun, would you like to try a probability example which is a little harder but which may especially fascinate your friends? Have you ever wondered what the chances are that two of your friends have the same birthdate? How many friends do you think you'd have to check with for the chances to be better than even that this will be the case? Are you surprised to hear that this number need be only 24? Let's prove this by using what we know about probability.

How do we approach this problem? We use a very helpful and yet simple method which scientists and mathematicians often employ: We merely turn the question around. Instead of asking whether such and such is true, we ask if the opposite question is true. In this case we start by finding the probability that each of the 24 people has a **different** birthdate.

To illustrate: The first person we ask can have any birthdate out of the 365 possible. Then the chances of the second person having a different birthdate are $\frac{365-1}{365}$, or $\frac{364}{365}$, since the first person used the first date ($365 - 1 = 364$). The probability that the third person has a different birthdate is still 1 less, or $\frac{363}{365}$. What is the probability for the fourth person? Yes, $\frac{362}{365}$. These probabilities continue along this line until the 24th person is reached. His probability of having a different birthdate is $\frac{365-23}{365}$ or $\frac{342}{365}$.

The probability that no two of these 24 people have the same birthdate is the **product** of all these separate probabilities, as shown below:

$$\frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \frac{361}{365} \times \ldots \times \frac{342}{365}$$

(The three dots [...] stand for the probabilities from $\frac{360}{365}$ to $\frac{343}{365}$, inclusive)

We have not shown the work, but the product of these multiplications is 0.46. Remember that 0.50 (or .50) means that the chances are even. Therefore the chances are 46 in 100, that is, less than even, that no two of the 24 people have the same birthdate. On the other hand, the chances are 54 out of 100 that two or more of them **do** have the same birthdate! Would you have thought that this was true?

Test Your Knowledge of Probability

Have you enjoyed your journey through the world of chance? We hope so, for you are about to enter another interesting region of mathematics — the study of statistics. First, though, check your learning by doing the examples which follow. Check your thinking on page 40.

SET L

1. Which of the following statements is true?
   a. Statisticians are never concerned with the future.
   b. What has happened can sometimes help us to estimate what will happen.
2. What must we be careful to do when we use statistics to make predictions about the future?
3. How many events can occur when an ideal coin is tossed?
4. If either of two events can occur but both cannot occur at the same time, how does a mathematician describe them?

SET M

Use this table of weather forecasts and actual weather for one week to answer the questions which follow.

Day	Weather Forecast	Actual Weather
Sunday	Clear and hot	Clear and hot
Monday	Increasing cloudiness	Clear and warm
Tuesday	Cloudy	Cloudy
Wednesday	Showers	Showers
Thursday	Heavy rain	Cloudy but no rain
Friday	Clear and warm	Clear and warm
Saturday	Sunny and hot	Sunny and hot

1. How many of the weather forecasts were correct? How many were incorrect?
2. Using this table as a model, what is the probability that this weather forecaster will be right next Friday?
3. Why are weather forecasts of "clear" and "cloudy" mutually exclusive events?
4. What is the probability of rain on a clear day?

SET N

There are three marbles in a box. One is red, one is white, and one is blue. The model below lists all the possible orders in which they can be removed from the box. Use this model to answer the following questions.

Red White Blue
Red Blue White
Blue Red White
Blue White Red
White Blue Red
White Red Blue

1. What is the probability that the first marble will be red?
2. What is the probability that the first marble will be red or white but not blue?
3. If the first marble is blue, what is the probability that the second marble will be white?
4. If the first marble is white, what is the probability that the second marble will also be white?
5. What is the probability that the first marble will be blue, the second red and the third white?

SET O

1. A drawer contains two black socks and two white socks. What is the probability that a blind man would remove a black sock first?
2. If two ideal coins are tossed, what is the probability that one will be heads and one will be tails?
3. If two pennies are tossed, what is the probability that neither will be heads?
4. If a penny, a nickel, and a dime are tossed, what is the probability that the penny will be heads?
5. How many different events can occur when three coins are tossed?

CHECKING YOUR THINKING

SET A

1. If one of the events occurs, the other cannot occur on the same toss.
2. Events are equally likely (chances are even) when there is no reason to expect one event to occur rather than the other.
3. Statistical probability
4. Mathematical probability
5. a. Mathematical. Remember, we said it was unnecessary to experiment when the situation was strictly defined. This is the case with the deck of cards, which consists of 52 cards, 4 of which are aces.
   b. Statistical
   c. Statistical
   d. Statistical

SET B

1. 1 (or 1.00)
2. 0 (or .00)
3. a. 0 (or .00)
   b. 1 (or 1.00)
4. Your list must consist only of events which are **certain** to happen such as:
   A man will die without oxygen.
   Water will boil at 212° F. at sea level.
   You are now looking at this page.
   Your mother is older than you.
5. Your list must consist only of events which cannot happen, which are impossible, such as:
   The stars will all fall down tomorrow morning.
   Mr. Brown will be in the state of New York while he is in the state of California.
   A dog can become a cat.
6. 0 (or .00)
7. 1 (or 1.00)

SET C

1 - 8. Answers for these examples will vary according to each pupil's tosses. Keep in mind that when you tossed your coin ideal conditions were not present. Remember too that even if you had ideal conditions and made 100, 1,000, 100,000, or 1,000,000 tosses, your probability would only approximate or approach the mathematical probability for coin tossing because you were gathering facts. You were using them to get statistical probabilities.

9. 1 (or 1.00)

SET D

1. $\frac{1}{2}$ or .50

2. $\frac{1}{6}$ or .166 or .17

3. **a.** $\frac{1}{6}$ or .166 or .17
b. $\frac{1}{6}$ or .166 or .17

4. **a.** $\frac{2}{6}$ or or .33
b. $\frac{2}{6}$ or or .33
c. $\frac{3}{6}$ or or .50
d. 1 (or 1.00)
e. 0 (or .00)

SET E

1. $\frac{7}{8}$

2. **a.** $\frac{1}{4}$
b. $\frac{1}{4}$
c. $\frac{3}{4}$

3. $\frac{1}{8}$

SET F

1. 16

2. $\frac{1}{16}$

3. $\frac{11}{16}$

4. $\frac{5}{16}$

5. $\frac{6}{16}$ or $\frac{3}{8}$

SET G

1. 2 and 12

2. $\frac{6}{36}$ or $\frac{1}{6}$

3. $\frac{4}{36}$ or $\frac{1}{9}$

4. $\frac{6}{36}$ or $\frac{1}{6}$

5. $\frac{30}{36}$ or $\frac{5}{6}$

6. $\frac{18}{36}$ or $\frac{1}{2}$

7. $\frac{18}{36}$ or $\frac{1}{2}$

8. $\frac{24}{36}$ or $\frac{2}{3}$

9. $\frac{13}{36}$

10. $\frac{2}{36}$ or $\frac{1}{18}$

11. $\frac{8}{36}$ or $\frac{2}{9}$

12. $\frac{28}{36}$ or $\frac{7}{9}$

13. 0 or .00

14. $\frac{9}{36}$ or $\frac{1}{4}$

15. $\frac{6}{36}$ or $\frac{1}{6}$

SET H

1. 6
2. 27
3. There are 6 spaces to fill. The first one can be filled with any one of 3 letters, the second can also be filled by any one of 3 letters, and so can the third. The fourth, fifth, and sixth places can be filled in 3 different ways. Therefore, by multiplying, you get: $3 \times 3 \times 3 \times 3 \times 3 \times 3 = 729$.
4. $6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$
5. There are 6 spaces to fill. The first one can be filled with any one of the 24 letters, as can the second and the third spaces. The fourth, fifth, and sixth spaces can each be filled in 10 different ways. Therefore, by multiplying you get $24 \times 24 \times 24 \times 10 \times 10 \times 10 = 13{,}824{,}000$.

SET I

1. a. $\frac{1}{3}$ or .33
   b. $\frac{2}{3}$ or .66 or .67
2. a. $\frac{3}{7}$ or .43
   b. $\frac{4}{7}$ or .57
3. a. 3
   b. $\frac{1}{2}$ or .50
   c. $\frac{1}{2}$ or .50
4. a. $\frac{2}{5}$ or .40
   b. $\frac{3}{5}$ or .60

SET J

1. a. $\frac{1}{2}$ or .50
   b. $\frac{1}{2}$ or .50
2. $\frac{1}{4}$ or .25
3. $\frac{1}{4}$ or .25
4. $\frac{1}{8}$ or .125 or .13
5. After removing 1 red and 1 white marble, you had 2 red and 2 white marbles left. The probability that your next selection would be red is $\frac{2}{4}$ or $\frac{1}{2}$.

SET K

1. a. $\frac{1}{4}$ or .25
   b. $\frac{3}{4}$ or .75
2. $\frac{1}{3}$ or .33
3. $\frac{1}{32}$ or .03

SET L

1. What has happened can sometimes help us to estimate what will happen.
2. We must evaluate the data to be sure that our facts are pertinent to the situation. Also, we must usually collect many facts in order to make dependable predictions.
3. Two
4. Mutually exclusive

SET M

1. 5, 2
2. $\frac{5}{7}$
3. If one occurs the other cannot occur at the same time.
4. 0 or .00

SET N

1. $\frac{1}{3}$
2. $\frac{2}{3}$
3. $\frac{1}{2}$
4. 0 or .00
5. $\frac{1}{6}$

SET O

1. $\frac{1}{2}$
2. $\frac{2}{4}$ or $\frac{1}{2}$
3. $\frac{1}{4}$
4. $\frac{1}{2}$
5. 8

You have learned to use some of the sharp tools of modern mathematics to cut through the seeming mystery surrounding chance, luck, and coincidence. You have seen how we can depend upon probabilities determined for given situations. No doubt you have concluded that though this is a world of chance, science and reason give us the understanding to live confidently within it.

Probability, as you know, tells us how often we can expect something to happen. It helps us to look into the future. But to make the most of what probability can tell us we need a companion science — statistics. Statistics is the science that enables us to get the facts in a form from which we can determine probabilities of complex events and predict future events.

In Part II of your booklet you will have the opportunity to learn about statistics and its uses. Remember to use the same methods of study you employed in Part I. Be sure that you understand each section before you start the section which follows. Do all the examples in each set; then refer to the answers. Ask for help from your teacher if you really need it, but do your best to reason through your difficulties on your own.

The Meaning of Statistics

In the first part of this booklet we referred briefly to statistics. Now we are going to take a closer look. The word **statistics** may not be too familiar to you, but what it stands for is very important in your life. Without statistics, your school district would not be able to decide how many new schools are needed or where they should be located. Without statistics, your state and local governments would not know where to build roads, how much fire and police protection to provide, or how to decide about many other essential services. Without statistics, scientists and scholars could not organize and interpret facts in ways which bring about so many changes in people's ways of living.

What does **statistics** mean? We might say that statistics are facts collected and put together or classified. Why is it necessary to put facts together and classify them? Think about the following situation. What if we made a list of all the children in your city's schools, showing each child's name, grade, teacher, and school, but in no regular order or pattern.Could you answer questions readily by reading such a list? No, not without a lot of searching.

But suppose your list classifies all the children according to the school they attend, the grade they are in, and the teacher they have. From this list you could quickly tell how many children are in each school, in each grade, and with which teacher. You would have a list with the facts organized.

Sometimes we use **statistics** as a singular word to mean the science or mathematical study of collecting and classifying facts so that we can grasp the important ideas and relationships that enable us to answer questions, solve problems, or make decisions. We call a person who collects, sorts, and classifies facts in this way a **statistician.**

The facts with which statistics deal are usually called **data, cases,** or **items,** or sometimes **members of a set.** In this booklet you will be introduced to some of the methods by which data, cases, or items can be organized and presented to help you see the overall picture or essential meaning of a whole collection of facts.

Some Ways of Organizing Facts

Tables and Graphs in Statistics

Find the words **table** and **graph** in the glossary in the back of this booklet. Have you seen or used a table or graph recently? Can you think of some ways you could use them in some project of your own? Perhaps you belong to a group which is raising money for summer camp, for the foster children program overseas, or for some similar purpose. Or maybe you are playing games and want to keep records of your scores over a period of time. Try to think of how a table or a graph could help you show your progress toward your goals.

Let's look at how tables and graphs help us to organize and present facts in useful ways. For our example, we might consider the facts that must be known and properly evaluated in order to have a well-managed school cafeteria. One kind of information that is helpful is the number of lunches served each day. This information is presented below in three different ways. As you work with each way, think about its advantages and disadvantages. Even though some questions are repeated, find the information asked for each time.

FIRST PRESENTATION

A Written Description

The numbers of children who were served by the school cafeteria for 10 days are as follows: April 1, 230; April 2, 320; April 3, 293; April 4, 187; April 5, 331; April 8, 284; April 9, 312; April 10, 256; April 11, 310; April 12, 80.

Use the description to answer the following questions.

SET P

1. On which date were the most children served? Which date was second highest?
2. On how many days was the number served more than 280?
3. Compare the number of children served in the period April 1-5 to the number served April 8-12 by careful examination (inspection) of the data. During which period, do you

SET P (cont'd.)

think, were more children served? (When we write dates [or other series of numbers] like this, April 1-5, we are indicating that we are including 1 and 5 as well as the numbers in between. The expression can be read: "April first through fifth, inclusive.")

4. What was the total number of children served during the 10 days?

Check your answers by turning to page 64.

SECOND PRESENTATION

A Tabular Presentation

Use Table I to answer the questions which follow it.

TABLE I

NUMBER OF CHILDREN SERVED IN JEFFERSON SCHOOL CAFETERIA FOR TEN DAYS

Date (Five-day week starting Monday)	Total Number of Children Served
April 1	230
2	320
3	293
4	187
5	331
8	284
9	312
10	256
11	310
12	80

SET Q

1. On which date was the smallest number of children served by the school cafeteria? On which date was next to the smallest number served?
2. On how many days was the number served less than 280?
3. Compare the number of children served April 1-5 to the number served April 8-12. During which period were more children served?

4. What was the total number served during the 10 days?
5. List some ways Table I would be useful to the school cook, the cafeteria manager, the principal, and the custodian.
6. Can you tell how many children were served each school day for the month of April by referring to Table I? Can you tell how many children of a certain grade were served each day during this month?
7. Can you take the information in Table I and put it together or summarize it? If so, how?

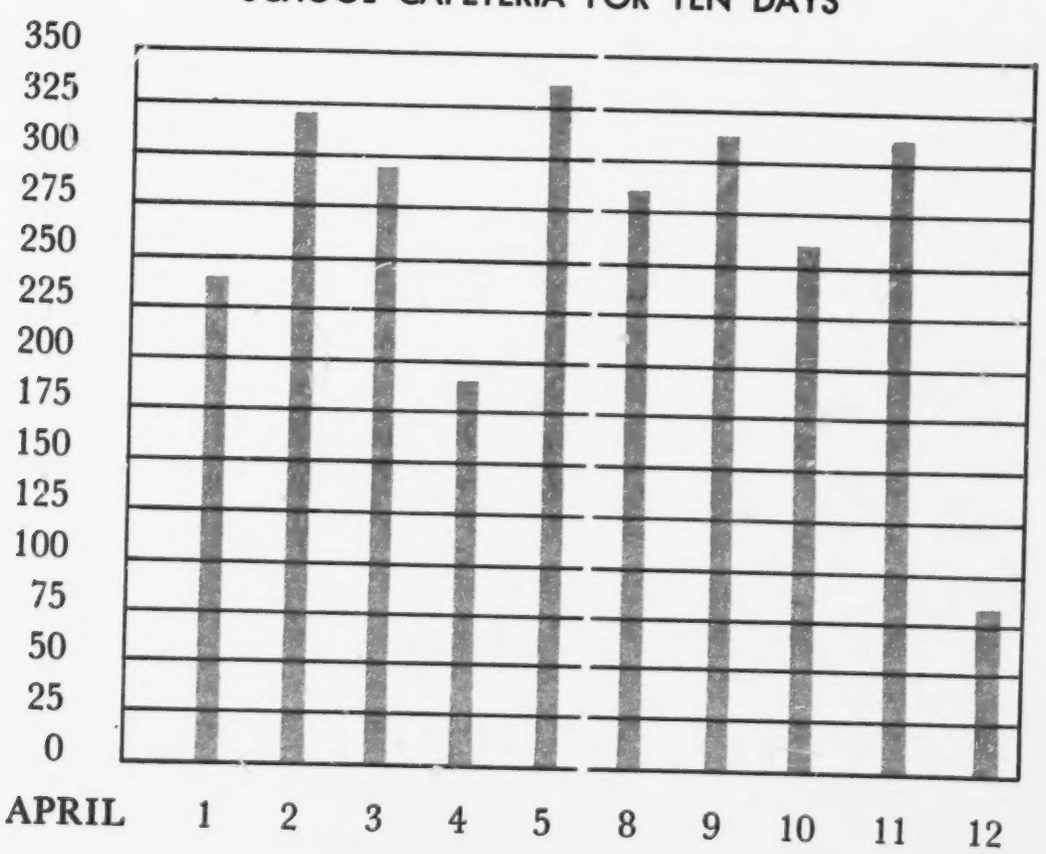
Check your answers by turning to page 64.

THIRD PRESENTATION

A Graphic Presentation

The information from Table I is shown as a graph in Figure I below. Statisticians often use graphs to show relationships among facts. Study this graph and use it to answer the questions in SET R on page 46.

FIGURE I. NUMBER OF CHILDREN SERVED IN JEFFERSON SCHOOL CAFETERIA FOR TEN DAYS



SET R

1. On which date were the most children served? The least?
2. On how many days was the number served less than 250?
3. Compare the number of children served in the period April 1-5 to the number served April 8-12 by looking at the graph. During which period, do you think, were more children served?
4. Can you tell from the graph the exact number of children served during the ten days?

Turn to page 64 to check your answers.

Now, for answering the types of questions which were asked, which way of presenting the data seemed easiest to use?

Have you remembered to think about the advantages and disadvantages of these three different ways of presenting data? Are you ready to make some decisions about them? See what you know and how you feel about these methods by answering the questions in SET S.

SET S

1. Which method of presenting data would take the most time to prepare: the written description, the table, or the graph? Why?
2. Does the graph help you to see the **general** differences in the numbers of children served more clearly than the table does? In what way?
3. Would it be better to use the graph or the table to find the exact number of children served by the cafeteria for a given date or week, or could either be used? Why?
4. Which way would be faster to find the date on which the most children were served—the table or the graph? Why?
5. Would the graph have been as effective if the number of children served had ranged from 275 to 325 instead of from 80 to 331? Why or why not?
6. If you wanted an overall idea at a glance, would you rather use the table or the graph? Why?

Check your work by turning to page 65.

In Table II we have included the information given in Table I together with some additional data. Use Table II to answer the questions in SET T.

TABLE II

NUMBER OF CHILDREN IN EACH GRADE SERVED IN JEFFERSON SCHOOL CAFETERIA FOR TEN DAYS

Date	Total Served	Grade 1	2	3	4	5	6
April 1	230	50	40	46	35	36	23
2	320	58	64	70	40	56	32
3	293	60	50	53	41	42	47
4	187	54	43	50	40	0	0
5	331	68	65	49	48	53	48
8	284	71	51	44	41	38	39
9	312	61	52	52	49	50	48
10	256	64	51	42	36	32	31
11	310	60	62	54	48	55	31
12	80	12	10	10	15	17	16

SET T

1. On which date were the most children served for any one grade? What was that grade?
2. What does Table II show you about which children were served on April 4?
3. What does Table II show you about how the data were distributed on April 12? The term **distributed** means the way in which the total number of cases were spread among the six grades.
4. Compare the numbers of children in each grade served on April 4 and April 12. What might have happened on these days to change the pattern?
5. By using the data in Table II, how can you find out which grade had the most children served for the period shown? Which grade was this?
6. To whom would such information as that found in Table II be valuable?

Check your answers to turning to pages 65 and 66.

Measures of Central Tendency

The example you have just worked through showed how you can arrange facts in a systematic way so that they are easier to study and understand. Using tables and a graph you studied such facts as the total number of children served in the school cafeteria for a ten-day period, the date when the most children were served, the date when the fewest children were served, and so on. These were findings about the past. From such findings you might attempt to predict the probability of future events. For example, you might try to forecast the dates which would be busiest for the school cafeteria next year if a similar pattern should be followed.

This is the kind of work the statistician does. He tries to determine what has happened in the past by studying the data he has gathered; then he tries to predict what may happen in the future. He examines the data carefully to determine what patterns exist, and whether changes or trends are appearing. In addition to the simple tables and graph which we have been using, the statistician has developed several efficient tools to help him in his work. Why, do you think, does he need these additional tools?

Suppose that instead of having just one table or graph showing the numbers of children served during a ten-day period, you had 18 such tables or graphs covering a 180-day period. If you needed to describe this data and the relationships that were important, how could you do it? First of all, you would have to summarize your 18 separate sets of data, wouldn't you? Further, you would want to have some way of showing the relationships within the data.

As you study this section of your booklet you will learn how the statistician summarizes a set of data and measures its main or typical characteristics. He usually tries to find a single numerical value that will represent all the data in a definite way. To see how he does this, let's take a look at how a set of data is spread or distributed.

In many large groups of data the cases or items tend to bunch or cluster about some central value. Usually, but not always, this bunching together occurs near the middle of the

distribution, which is why we use the term **central tendency.** Statisticians have developed several ways of measuring central tendency. We will examine three of these. Each of the three has advantages for some purposes and disadvantages for others. As we think through an example about a summer camp building fund, try to see what the different measures tell us and why they are not equally useful in all situations.

The Arithmetic Mean

Suppose you want to find the average of the set of numbers 5, 6, 8, 9, and 7. Look at the box below to follow the steps. First you add the numbers to get their sum (35). Second, you count to get the number of items (5). Finally, you divide the sum by the number of items in the set.

Step 1 Add to get sum	Step 2 Count the items	Step 3 Divide to get arithmetic mean
5	1	
6	1	
8	1	$5\overline{)35}$
9	1	
7	1	
35	5	

The quotient, 7, is the average of the data. Statisticians use this method to find the average, but they have a special name for it. They call it the **arithmetic mean,** or simply the **mean.** The formula statisticians use to show how they find the mean can be written in words as follows:

$$\text{Arithmetic mean} = \frac{\text{sum of numerical values of all members in the set}}{\text{number of members in the set}}$$

See how you can use this formula as you work on our camp building fund example, which follows.

A group of people are planning to build a camp for boys and girls. They have asked for contributions for the building fund. At this point 17 people have contributed the following amounts (in dollars): 5, 10, 10, 5, 15, 25, 10, 20, 15, 30, 25, 15, 25, 15, 15, 10, and 5. What is the **mean** of these contributions?

To find out, what do you do first? Yes, add the numerical values of all the contributions. What sum do you get? $255.00. Now, what is the next step? Yes, divide by the number of items in the set. How many items are there? There are 17 contributions. What is the average or mean contribution? $255 \div 17 = 15$. The average or mean contribution is $15.00.

Now try your hand at being a statistician. Use what you have just learned to do the work in SET U.

SET U

Table III shows many people visited United States national parks during a 10-year period. Use this table to answer the questions which follow it.

TABLE III

NUMBERS OF PEOPLE VISITING UNITED STATES NATIONAL PARKS DURING A TEN-YEAR PERIOD

Year	Number of visitors (in millions)	Yearly Increase (in millions)
1946	21.8	—
1947	25.5	3.7
1948	29.9	4.4
1949	31.7	1.8
1950	33.3	1.6
1951	37.1	3.8
1952	42.3	5.2
1953	46.2	3.9
1954	47.8	1.6
1955	50.0	2.2

1. a. What was the total number of visitors to national parks for the period from 1946 through 1955?
   b. What was the mean number of visitors per year to national parks for the period shown?
   c. What was the mean yearly increase in the number of visitors per year to the nearest tenth of a million?

Check your answers by turning to page 66.

Table IV below shows the population counts in millions for the United States during the period from 1910 through 1960. Use this table to answer the questions which follow it.

TABLE IV

POPULATION OF UNITED STATES FOR TEN-YEAR PERIODS 1910 - 1960

Year	Population in Millions	Increase in Millions
1910	92.0	—
1920	105.7	13.7
1930	122.8	17.1
1940	131.7	8.9
1950	150.7	19.0
1960	179.3	28.6

2. **a.** What was the mean population to the nearest tenth of a million from 1910 through 1960?

b. What was the total increase in population from 1910 through 1960?

c. What was the mean increase in population per decade (10-year period) from 1910 through 1960?

d. During which of these decades did the population show the smallest increase?

e. During which of these decades did the United States population show the greatest increase?

Check your answers by turning to page 66.

3. See Table II on page 47 to answer the following questions.

a. How would you find the mean number of children served per grade on any date?

b. What was the mean number of children served per grade on April 9? Did any grade or grades actually have this number served that day?

c. How would you find the mean number of children served per day for any one grade during the period shown?

d. What was the mean number of children served in the third grade for the period of time shown?

e. Find the mean of the total number of children served in the period of time shown. (Round your answer to the nearest whole number.)

Check your answers by turning to page 66.

The Median

The mean is a measure of central tendency very useful in handling certain groupings of data, but not so useful when a few very large or very small cases are included. Often we want a measure that will tell us what the middle number is in a set of data if the data are listed in order from largest to smallest. The kind of measure that does this is called the **median.**

Let's look at our camp building fund again. The amounts given were (in dollars): 5, 10, 10, 5, 15, 25, 10, 20, 15, 30, 25, 15, 25, 15, 15, 10, and 5. You found that the **mean** of these amounts was $15.00. How would you find the **median** or middle amount contributed?

30	1	17
25	2	16
25	3	15
25	4	14
20	5	13
15	6	12
15	7	11
15	8	10
15	9	9
15	10	8
10	11	7
10	12	6
10	13	5
10	14	4
5	15	3
5	16	2
5	17	1

First you must arrange these numbers in some order. You could arrange the data from lowest to highest, but it is customary to arrange data from highest to lowest, as we have done here. This is called ranking the data.

How do you find the middle or median contribution? If you look at the table at the left, you can see that by counting down 8 items you reach the middle case. The ninth case is the middle case because there are just as many cases above it (8) as there are below it. You can check this if you wish by counting up from the bottom. The median figure is $15.00. How does this compare with the mean? Yes, it is the same.

However, suppose that the camp building fund has now grown to $3465.00 with the addition of 4 more people contributing. Three of these people gave $1000.00 each, the fourth gave $210.00. We know what the original 17 contributors gave. Now, would the mean or the median give the better measure of central tendency? Let's see if we can figure this out.

You can compute the mean by dividing the total by the number of contributions. 3465 ÷ 21 = 165. $165.00 is the mean contribution now. You can rank your contributions once more.

1000
1000
1000
1000
210
30
25
25
25
20
15
15
15
15
15
10
10
10
10
5
5
5

When you count down or up 10 cases you find that the median contribution is still $15.00. Look at the ranked data once more. Which measure of central tendency gives us a better picture of the data — the mean, $165.00, or the median, $15.00? In other words, what number of dollars is more typical of **all** the gifts? Are more of the contributions like $165.00 or are more like $15.00? Yes, more are like $15.00, so the median gives us a more typical measure of the central tendency. The mean, $165.00, is not very close to either the largest contributions or the smaller ones. Only four gifts are more than $165.00. Three of these are a great deal more, so the mean doesn't really describe them at all. Seventeen of the gifts are very much less than the mean. In fact, only one gift ($210.00) is at all close to the mean. On the other hand, there are many gifts not too much greater or smaller than $15.00. In our example, then, $15.00 is a much more typical gift than $165.00. Thus you can see how a median can be much more useful than a mean in describing some sets of data.

In our camp building fund examples we always had an uneven number of cases, so the middle case stood out clearly once we ranked the data and counted. Suppose now you need to find the median when there is an **even** number of cases in a set of data. Look at this list of the weights of six boys: 92, 90, 89, 85, 83, and 82 pounds. Now, what is the median weight? If you rank the data you can count down 3 weights and come to the 89-pound case, while if you count up 3 weights you stop at the 85-pound case. There is no actual middle case, but since you are describing the typical characteristic of the whole set of

data, you can find the mid-point between the two cases closest to the center and use this as the median. One way to do this is to add the weights and divide by 2. First add: $89 + 85 = 174$. Then divide: $174 \div 2 = 87$. The median is 87 pounds. Another way is to find the difference between the two weights, divide this by 2, and add the result to the lower case or subtract it from the upper case.

When will you want to use a median instead of a mean? Can you think of times when a mean might be more helpful than a median? Keep these questions in mind as you do the examples in SET V.

SET V

1. Find the median of these test scores: 85, 75, 79, 62, 64.
2. Find the mean of the test scores in Example 1. How does it compare with the median for this set of data?
3. Find the median weight of 9 boys whose weights are 68, 72, 85, 51, 77, 73, 65, 96, and 79 pounds.
4. Find the mean weight of the boys in Example 3. How does the mean compare with the median in this case?
5. In Example 3, would the median be changed if the boy weighing 51 pounds were replaced by a boy weighing 69 pounds? Would the mean be changed? Prove your answers by finding the median and the mean after substituting the weight of the boy weighing 69 pounds for the weight of the boy weighing 51 pounds.
6. Fifty people gave \$3450.00 to a United Fund drive. Two people gave \$1000.00 each, the other forty-eight people's gifts were from \$5.00 to \$60.00 apiece. Would the mean or the median give the better measure of central tendency? How can you prove your answer?
7. The following numbers tell how many copies of a certain book were sold for each of 10 days: 50, 47, 45, 42, 35, 35, 35, 34, 34, 33. Would the mean or the median give you a better measure of central tendency? Why?

8. Find the median of this set of numbers: 24, 39, 65, 75, 61, 45.
9. Find the median of this set of numbers: 3, 7, 9, 11, 13, 15.
10. Find the median height of 8 girls whose heights are 51, 58, 64, 60, 63, 59, 60 and 65 inches.
11. What is the mean height of the eight girls in Example 10? How does the mean compare with the median?

Check your answers by turning to page 67.

The Mode

A third measure of central tendency used by statisticians is called the **mode.** The mode is the case or item which occurs most frequently in a set of data. We find the mode by inspection, that is, by scanning the data after it is ranked to see which case or item is listed the most times.

You can do this yourself by referring to our list on page 53 of twenty-one contributions to the camp building fund. Which contribution occurs the most times? Yes, $15.00, which occurs 5 times. We call $15.00 the mode of this set of data, considering it the most typical of all the cases. Note that this is the same as the median in this particular distribution, and that the mean varies widely from both the mode and the median.

YELLOWTAIL WEIGHTS

28 lbs.
24
24
24
22
20
19
19
18
198

Now look at the list of weights at the left. The list was made from the weigh-in of 9 yellowtail brought in by a sportfishing boat. You can see by inspection that the mode is 24 pounds. What is the median? Yes, the median is 22 pounds. The mean, 198 pounds divided by the number of fish, is also 22 pounds. This situation, then, is different from our previous one. Now the mode is higher than the mean and the median, which happen to be the same.

You can see that while the mean, the median, and the mode are all called "measures of central tendency," they may or not be the same number in a given set of data.

Businessmen often use the mode to find the most popular styles or colors of the things they sell or manufacture.

NEW CAR SALES	
Color	Number Sold
Red	3
White	8
Blue	5
Green	6
Black	1
Tan	2
Yellow	4
Grey	8

The table at the left shows the number of new cars of certain colors sold by one auto dealer. What is the largest number of sales of any one color? Yes, 8. Can there be more than one mode? In our example both white and grey appear 8 times, so there are two modes. We say this set of statistics is **bimodal.** Can you figure out why it is called this?

Now do the examples in SET W. Keep in mind all three measures of central tendency we have discussed.

SET W

1. Which of the measures of central tendency you have studied would the following people be most likely to use?
   a. A statistician who wishes to know the average weight of 11-year-old boys.
   b. A dress manufacturer who wishes to know what color is most popular.
   c. An employer who wishes to know the average weekly salary of the workers he employs.
   d. A teacher who wishes to know the number of words half of the children in her class can spell correctly on a test.
2. Is it possible for the mean, the median, and the mode all to be the same value?
3. What is the mode for the set of data in Example 7, SET V?
4. If a chicken rancher tells you he gets an average of 2275 eggs a week from 500 laying hens, he has given you the measure of central tendency called the_____?
5. If the chicken rancher tells you that half of his hens give more than 5 eggs a week while the others give less, which measure of central tendency is he using?
6. If the rancher says that more of his hens lay 5 eggs a week than any other amount, he has given you which measure of central tendency?

Check your answers by turning to page 67.

The Range

One of the important features of any set of data is its **range.** This is very easy to find because it is simply the difference between the highest and the lowest numbers in the set. For example, to find the range for Table I on page 44 we subtract 80 (the lowest number of children served lunches) from 331 (the highest number of children served lunches) and we find that the difference, or range, is 251. Since this range is a very large number, we can say that there is a considerable difference in the number of children served lunches in the cafeteria for the various school days of this 10-day period. The range, then, is a measure of the total spread of the data. It tells how widely the set of data is spread out from largest to smallest or from end to end. We can say that the range measures the total difference among the data.

Look at Table III on page 50 again. How would a statistician use the range to describe this data? Use Table III to work Examples 1-5 in SET X below. Check your thinking on page 68.

SET X

1. What was the range in the number of visitors for the period from 1950 through 1955?
2. Find the range of the increase in the number of visitors as shown in the third column of Table III.
3. In which year during the period from 1948 to 1955 did the number of visitors show the greatest increase over the previous year?
4. During which of the years 1948-1955 did the number of visitors show the smallest increase?
5. About how many times as many visitors were there in 1955 as in 1947?
6. What was the range of the population in Table IV, page 51?
7. What was the range of the data ranked on page 53?

Grouping Statistical Data

Often there are so many members in a set of data that it is not convenient to list each member separately. At the left is a set of test scores made by 15 pupils. These scores are listed in rank order from highest to lowest. What is the mean for this set of data? The median? The mode? The range? (Turn to page 68 to check your answers.) Listing each score separately uses more time and space than is necessary when several scores appear more than once, as they do in our list here. Statisticians have found a way to list each score only once by using a second column to show how many times each score appears in the set of data. They call this second column the "frequency" column because the word **frequency** is a shorter way of saying "number of times a case or item appears." They would write the set of scores as shown in Table V below.

TEST SCORES
67
65
64
63
63
62
62
62
61
61
60
59
59
57
56
921

TABLE V

FREQUENCY DISTRIBUTION OF TEST SCORES OF FIFTEEN STUDENTS

Scores	Frequency	Products of Frequency times Score
67	1	67
65	1	65
64	1	64
63	2	126
62	3	186
61	2	122
60	1	60
59	2	118
57	1	57
56	1	56
Total frequency =	15	Total of scores = 921

Statisticians call this a "Frequency Table." The third column is necessary because we can find the mean only when we know the sum of all the scores in the set of data. When data are grouped as in Table V each score is written only once. To find the sum of all the scores, the statistician multiplies each score by its frequency and then finds the sum of these products.

Is the sum of the products the same as the sum of all the scores listed separately? Check back to see. The statistician now finds the number of scores by adding the numbers in the frequency column. Is the sum the same as it was when we counted each score separately? Dividing 921 by 15, we find that the mean is 61.4, which is the same mean we found before these scores were grouped.

You have learned that the median for ranked data is the middle number or case. How many test scores are there all together in the list on page 58. How far up will we have to count to find the middle score? How far down? The eighth number is what score? Yes, 62. How, do you think, can we find the median using Table V? Try counting up or down in the frequency column to see where the eighth-ranked case is. Then look in the first column to find the value of this group of scores. As you can see, the median in our grouped data is also 62 since the eighth number is one of the three scores of 62.

It is easy to find the mode in a frequency table. Since it is the number that appears most frequently, we simply look for the largest number in the frequency column. In the set of data in Table V, what is the largest frequency for any score? What is that score? (Check your answers on page 68.) Is the mode the same as when the scores were listed separately?

Finding the range from a frequency table is also simple. Since it is the difference between the highest and the lowest items in a set of data, what do we do to find it? Yes, subtract. 67 minus 56 is how many? Is the range of 11 the same as it was before we made the frequency table?

Think carefully as you do SET Y which follows. You will be using almost everything you have learned so far in Part II of your booklet. Be sure to review any section you need to.

SET Y

1. Make a frequency table for the following set of data which gives the ages of children on an all-school committee: 10, 15, 16, 16, 14, 10, 13, 15, 11, 12, 10, 14, 9, 15 ,11, 17. Check to be sure that the total of frequencies in your table is 16.
   a. Find the mean of the set of data above.
   b. Find the median of the set of data.
   c. Find the mode of the set of data.
   d. Find the range of the set of data.
2. Make a frequency table for the following number of miles of daily travel recorded for one car:
   14, 6, 7, 7, 13, 8, 13, 12, 8, 8, 10, 12, 9, 9, 10, 11, 12, 9, 9, 11, 10, 10, 11, 10, 11. Is 25 the total of the frequency column of your table?
   a. Find the mean of the set of data above.
   b. Find the median of the set of data.
   c. Find the mode of the set of data.
   d. What is the range of the set of data?

Check your answers by turning to page 68.

Grouping Data Into Step Intervals

When the range of a set of data is large, statisticians group the data by combining two or more of the cases or items into what they call a **step interval.** Suppose we measured the heights of 100 fourth grade pupils selected at random and found that the tallest pupil was 64 inches in height, the shortest was 40 inches, and the other 98 were spread out between these two heights. If we made a frequency table listing all the heights between 40 and 64 inches, it would be quite long. Statisticians usually prefer frequency tables that have no more than ten steps.

Since the range of this set of data is 24 inches, the statistician sees that by combining the heights into steps of 3 inches each he will need only 9 steps in his table. Since the tallest pupil is 64 inches in height, the highest step in the table will include the group of pupils who are 62, 63, and 64 inches tall—a step interval of 3. How many inches tall will the pupils be who are in the next highest step? See how the statistician sets up step intervals on the following frequency table.

TABLE VI

FREQUENCY DISTRIBUTION OF HEIGHTS OF 100 FOURTH GRADE PUPILS SELECTED AT RANDOM

Heights in inches	Frequency
62-64	4
59-61	7
56-58	12
53-55	17
50-52	20
47-49	17
44-46	12
41-43	7
38-40	4
Total Frequency =	100

From this frequency table we cannot find the exact median or mode in inches. Instead, we can only say that the median and the mode are between 50 and 52 inches.

Suppose we wish to find the mean height of the 100 pupils from this set of grouped data. Can we find it by using our table? No, we cannot, because the table does not give the individual heights. Therefore we must set up our data differently. Now look at Table VII. Notice that two new columns have been added.

TABLE VII

FREQUENCY DISTRIBUTION OF HEIGHTS OF 100 FOURTH GRADE PUPILS SELECTED AT RANDOM

Height in inches	Mid-Point of Step Interval	Frequency	Mid-Point Times Frequency
62-64	63	4	252
59-61	60	7	420
56-58	57	12	684
53-55	54	17	918
50-52	51	20	1020
47-49	48	17	816
44-46	45	12	540
41-43	42	7	294
38-40	39	4	156
	Total Frequency =	100	Total = 5100

Because we did not know the individual heights of the pupils in any of these nine steps, or groups, we made an assumption. We assumed that the four pupils who are 62-64 inches tall have an average or mean height of 63 inches since 63 is halfway between 62 and 64. Notice that this mean, 63, is really the median of the numbers 62, 63, and 64. We make the same assumption for each group of heights and list these mid-points, or medians, in the second column of the frequency table.

Now think: How do we get the data in the fourth column? Is it the same as the total height of each group of pupils? How do we get the total of the heights of these 100 pupils? What is the mean height of these 100 pupils? Did you divide 5100 by 100? Check your answers by turning to page 69.

Do you remember what the median and the mode are for this set of data? Like the mean (51 inches), they are also between 50 and 52 inches. So this is one set of data in which these three kinds of measures of central tendency are all about the same.

Test Your Knowledge of Statistics

You have studied some of the ways statisticians go about their work of collecting, classifying, and studying facts. You have checked on your learning as you went along by doing examples as each new step was developed. The self-evaluation examples in SET Z that follow will help you to evaluate what you have learned.

If you have trouble or need help, review what you have studied to see if you can find the answer to your question or problem by yourself.

SET Z

1. a. How would you define statistics?
   b. Must a statistician be a good mathematician?
   c. What does a statistician call the facts he collects?
   d. How do tables and graphs help in studying a set of data?
   e. List three measures of central tendency which we used in this booklet.

Use Table VIII on page 63 to answer the questions which follow it.

TABLE VIII

MEAN NORMAL DAILY TEMPERATURES FOR THREE CALIFORNIA CITIES

Month	Los Angeles	Sacramento	San Francisco
January	55.0	45.2	50.1
February	56.4	50.6	53.0
March	58.9	54.9	54.9
April	61.5	59.2	55.7
May	64.8	65.0	57.1
June	67.8	71.2	59.1
July	72.5	75.3	58.9
August	72.9	74.0	59.3
September	71.0	71.3	61.6
October	66.6	63.8	61.0
November	62.1	54.3	57.2
December	57.3	46.7	51.9

2. **a.** Find the mean annual temperature for each of the three cities.
   b. Find the median annual temperature for each of the three cities.
   c. Find the range of the temperatures for each of the three cities.
   d. During which month of the year is the mean normal daily temperature most nearly the same in all three of these cities?
   e. If you wanted to spend a cool summer, which of these cities would you visit?
3. Make a frequency table of the following points made by a basketball team during one season: 58, 57, 58, 54, 62, 61, 57, 55, 55, 60, 58, 56, 60, 59, 57, 58, 58, 57, 56, 59. Use this table to answer the following questions.
   a. What is the mode of this set of data?
   b. What is the median of the data?
   c. What is the mean of the data?
   d. What is the range of this set of data?

Check your answers by turning to page 69.

CHECKING YOUR THINKING

SET P

1. April 5. April 2.
2. 6
3. April 1-5
4. 2603

SET Q

1. April 12. April 4.
2. 4
3. April 1-5
4. 2603
5. It could be used in ordering food, arranging eating space, providing for supervision of pupils, planning the length of the lunch period, determining the amount of help needed in cafeteria and lunchroom. Can you think of other ways?
6. No. To get data for the number of children served each school day in April you would need to collect data for the entire month. To find the number of children of a certain grade who were served you would need to collect the information by grades rather than for the entire school. Can you see why you need to know for what purpose you are collecting data before you begin to collect it?
7. You could total the data by weeks and list the numbers of children served each week. Or you might total the number of children served by days of the week—how many on Mondays, how many on Tuesdays, etc.

SET R

1. April 5. April 12.
2. Three
3. It is difficult to tell this from looking at the graph. You would probably have to refer back to Table I to be sure of your answer.
4. No. The table is better for adding accurately.

SET S

1. The graph. The written description and the table require only that the data be presented in numerical form. The graph requires that the data be shown by means of lines or bars as well as by numbers and words.
2. Yes. The presentation of the data in the graph gives you a clear picture of the daily totals at once. You can see the variations in these amounts at a glance.
3. You would have to use the table to find the exact number of children served for a given day, or even for a week. The graph gives only an approximation of totals.
4. The graph. You can immediately pick the highest point (or bar) on the graph and see which date it stands for. Note that in using the table to find the largest number you have to scan the hundreds' column first, then scan again to combine the hundreds' and the tens' columns, and then scan the ones' column.
5. No, the range would be too small unless a large portion of the graph were left out to save space so that the spaces from 275 to 325 could be made larger. If the graph spacing were left just as it is, the tops of the bars would all be close together and the picture of any variations would not be very clear.
6. The graph. Its pictorial arrangement of the data lets you see the daily totals both separately and as a group so that you can make immediate comparisons and judgments. Relationships among the data are apparent at once.

SET T

1. April 8. Grade 1.
2. That about the usual number of children in Grades 1-4 were served but that none in Grades 5 and 6 were served.
3. That each grade had some children served by the cafeteria but that the total number and the number for each grade were quite small compared to the usual amounts.

SET T (cont'd.)

4. The breakdown of the data in Table II makes it possible for you to see that the situations on April 4 and 12 were not the same, since **no** children in grades 5 and 6 were served on April 4 whereas **some** children in **all** of the grades were served on April 12. An event such as a class picnic, field trip, or some special occasion for grades 5 and 6 might have taken place on April 4, while on April 12 the event might have been school-wide but not involving all of the children.
5. By adding the vertical columns to get the total served in each grade for the ten days. Grade 1.
6. The principal, the cafeteria manager, student helpers in the cafeteria, and so on, to help in planning food preparation, seating space, number of helpers required.

SET U

1. **a.** 365.6 million
   b. 36.56 million. Did you find that in any one year exactly 36.56 million people visited the national parks? Look back to Table III to see. No, there was no such figure for any year. However, if 36.56 million people had visited national parks in each of the years from 1946 through 1955, the total number of visitors would have been 36.56 million times 10, or 365.6 million. This is the same as the total number of visitors for the ten years.
   c. 3.1 million
2. **a.** 130.4 million
   b. 87.3 million
   c. 17.5 million
   d. 1930-1940. These were the depression years.
   e. 1950-1960
3. **a.** Divide the total number of children served for all the grades for that date by the number of grades served.
   b. 52. Yes, Grades 2 and 3.
   c. Total the number of children served for that grade for the period shown and divide by the number of days.
   d. 47
   e. 260

SET V

1. 75. If you thought the median was 79 what did you forget to do? Yes, you forget to rank the scores first.
2. 73. It is 2 less than the median.
3. 73 pounds
4. 74 pounds. It is larger than the median by 1 pound.
5. No, the median would not be changed. Yes, the mean would be. The median would still be 73 pounds. The mean would be 76 pounds.
6. The median would. First compute the mean by dividing $3,450.00 by 50 ($69.00). Then try to estimate the median. Subtract the two $1,000.00 donations from $3,450.00. The $1,450.00 remaining was given by 48 people, whose gifts ranged from $5.00 to $60.00. The median gift of all 50 cases is between the twenty-fifth and twenty-sixth cases. This is between $5.00 and $60.00 and would probably be somewhere around $30.00 if these remaining gifts were evenly distributed. In any case, the median couldn't be $69.00, the amount found to be the mean. Thus we see that the mean would not give us an accurate measure of the central tendency of this set of data. Note that we can tell this even though we do not actually have a complete listing of the items.
7. The mean. The mean is 39; the median is 35. If you were describing the central tendency of this set of data, which goes from 33 through 50, you would choose 39 as more typical of the whole set.
8. 53. Remember, when there are two middle scores you add them together and divide by 2.
9. 10
10. 60 inches
11. 60 inches. The mean and the median are the same value.

SET W

1. a. Mean
   b. Mode
   c. Mean
   d. Median
2. Yes
3. 35
4. Mean
5. Median
6. Mode

SET X

1. 16.7 million
2. 3.6 million
3. 1952
4. 1950 and 1954
5. About 2 times.
6. 87.3 million
7. 995 (dollars)

Test Scores

The mean is 61.4.
The median is 62.
The mode is 62.
The range is 11.

Table V Answers

3 for the score of 62
Yes, the mode is still the same.
Yes, the range is still the same.

SET Y

1. FREQUENCY TABLE

Ages	Frequency	Products
17	1	17
16	2	32
15	3	45
14	2	28
13	1	13
12	1	12
11	2	22
10	3	30
9	1	9
	16	208

a. 13 **b.** 13.5 **c.** 10 and 15 **d.** 8

2. FREQUENCY TABLE

Miles	Frequency	Products
14	1	14
13	2	26
12	3	36
11	4	44
10	5	50
9	4	36
8	3	24
7	2	14
6	1	6
	25	250

a. 10 **b.** 10 **c.** 10 **d.** 8

Table VII Answers

By multiplying the second column (each mid-point) by the third column (the frequency). This gives us the totals of the heights of each group of pupils.
By adding the numbers in the last column.

SET Z

1. a. A collection of facts put together in such a way that they help us to recognize important ideas or relationships and to make decisions or solve problems. Also, the science of collecting and classifying facts in ways that help us to see general truths or ideas.
   b. Yes, because most statistics are in the form of numbers.
   c. Data. Also cases, items, or members of a set.
   d. They make it easier to study a set of data by presenting it in an organized arrangement.
   e. Mean, median, mode.
2. a. Los Angeles, 63.9 degrees; Sacramento, 61 degrees; San Francisco, 56.7 degrees.
   b. Los Angeles, 63.5 degrees; Sacramento, 61.5 degrees; San Francisco, 57.2 degrees.
   c. Los Angeles, 17.9 degrees; Sacramento, 30.1 degrees; San Francisco, 11.5 degrees.
   d. March
   e. San Francisco

3. **FREQUENCY TABLE**

Scores	Frequency	Products
62	1	62
61	1	61
60	2	120
59	2	118
58	5	290
57	4	228
56	2	112
55	2	110
54	1	54
	20	1155

3. a. 58 b. 58 c. 57.8 or 57.75 d. 8

GLOSSARY

arithmetic mean The numerical value obtained through dividing the sum of the members of a set of data by the number of items in the set (commonly called simply the **mean**).

average In this booklet, the arithmetic mean. To statisticians, **average** can be used as a general term to mean all measures of central tendency based on group data, including not only the **arithmetic mean** but also the **median,** the **mode** and others we do not discuss.

bimodal A set of data which has two members which occur the same number of times but more often than any other cases in the set; thus, a set of data with two modes.

central tendency The main or typical characteristic of a set of data.

data A group of facts or statistics.

distribution The way in which data are arranged or fall.

event a happening; an occurrence; something which takes place or occurs.

given event The particular event with which we are concerned; the event out of two or several events which we are predicting or counting.

graph A picture or drawing which shows the relation between certain sets of numbers.

independent events Events are independent if the occurrence or non-occurrence of one is not affected by the occurrence or non-occurrence of the other(s).

mathematical probability Probability determined through reasoning; probability based on what we already know and not requiring experimentation.

mean (See arithmetic mean)

measure of central tendency A single number or value which is used to describe an entire set of data by showing the main or typical characteristic of the whole set. The **mean, median,** and **mode** are measures of central tendency.

median The middle case in a set of data distribution; the point that divides the group of data into two equal parts.

mode The case or value that occurs the most often in a set of data.

model A selection of certain features of a situation to represent relationships in which we are especially interested. An example is a map showing distances, directions, and locations which stand for the actual physical features of a geographical area.

mutually exclusive events Two events are called mutually exclusive if the occurrence of either of the events makes it impossible for the other event to occur.

occur To happen; to take place.

permutation The arrangement of a group or set of things in a definite order or line.

range The difference between the lowest and the highest data in a set.

rank To arrange in order of numerical size.

ratio The numerical value of one quantity divided by the numerical value of a second quantity.

statistical probability Probability which is determined through experimenting, observing, and recording facts and events.

statistician A person who collects, sorts, and studies facts or data.

statistics Numerical facts collected and put together or classified. Also, the science of collecting and classifying facts in such a way that we can grasp important ideas or relationships.

step interval An equally divided distance between points or data in a set.

table A collection or arrangement of numbers or data in an orderly form.

This Independent Learning Booklet is part of an experimental project to provide mathematics materials for able pupils and to develop a mathematics program adapted to individual differences. The project is based on several studies and exploratory programs carried on in San Diego County during the past several years, such as those sponsored by the San Diego Metropolitan Education Study Council and the Assistant Superintendents Group. Objectives of this phase of the project are to stimulate and widen pupil interest in various areas of mathematics; to develop deeper insights into mathematical principles, concepts, and relationships; to introduce aspects of modern mathematical thinking; and to develop motivation for able pupils to continue to study and explore mathematics and related fields.

The project is coordinated by the Department of Education, San Diego County, and directed by a steering committee selected from members of the San Diego Metropolitan Education Study Council. Support for the project is derived from funds provided under Title III of the National Defense Education Act of 1958 and matching funds from the participating school districts and the San Diego County Department of Education.

BOARD OF EDUCATION
San Diego County

Martin W. Perry, *President*
Ray W. Booth
Edwin M. Campbell
P. N. Hyndman
Mrs. Mildred Macpherson
Dr. Cecil D. Hardesty, *Secretary*
Superintendent of Schools

Districts participating in the project are:

Alpine Union School District
Borrego School District
Cajon Valley Union School District
Cardiff School District
Carlsbad Union School District
Chula Vista City Schools District
Coronado Unified School District
Del Mar Union School District
Encinitas Union School District
Escondido Union School District
Julian Union School District
Lakeside Union School District
La Mesa-Spring Valley School District
Lemon Grove School District
National School District
Oceanside Union School District
Poway Union School District
Rancho Santa Fe School District
San Diego Unified School District
Santee School District
San Ysidro School District
Solana Beach School District
South Bay Union School District
Warner Union School District

Members of the Steering Committee are:

Dr. Paul E. Zintgraff, Chairman
Mrs. Leslie S. Beatty
Dr. George V. Hall
Dr. Richard Madden
Harry C. Rubie
Dr. Manfred H. Schrupp
John T. Warburton
Mrs. Edwena B. Moore, Project Director

Mathematics Consultant for this Independent Learning Booklet was:

Margaret F. Willerding

Reviewers for this booklet were:

Leslie S. Beatty
Charles B. Bell, Jr.
William E. Clarke
Robert T. Gray
Robert F. Heilbron
Edwena B. Moore
Joseph M. Moser

Other Independent Learning Booklets have been prepared and will be announced as soon as they are ready for distribution.